

Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 22, 2005, 11:22:19 ; Search time 163 Seconds
 (without alignments)
 1100.962 Million cell updates/sec

Title: US-09-719-379A-81

Perfect score: 2477

Sequence: 1 MDPTKLALSLLAAGVLAGCS.....YKPYDNKRIDSSTGHHHHHHH 464

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105632 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database :	A_Geneseq_16Dec04:*
1:	geneseqP1900B:*
2:	geneseqP1900B:*
3:	geneseqP2000B:*
4:	geneseqP2001B:*
5:	geneseqP2002B:*
6:	geneseqP2003B:*
7:	geneseqP2003B:*
8:	geneseqP2004B:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No. Score Query Match Length DB ID Description

Result No.	Score	Query	Match	Length	DB	ID	Description
1	2477	100.0	464	3	AAY79993		Aay79993 Plasmid L
2	1920	77.5	364	2	AAR20108		Aar20108 Protein D
3	1251.5	50.5	358	6	ABBB4614		Abba4614 E. coli 9
4	1219.5	50.5	359	7	ABD01848		Abd01848 Klebsiell
5	1210.5	49.2	364	7	ADF05718		Adf05718 Bacterial
6	1074	48.9	359	6	ABM68758		Abm68758 Photobact
7	661.5	43.4	356	2	AAY52773		Aay52773 Treponema
8	661.5	26.7	451	6	AAY06589		Aay06589 Lipoprote
9	651	26.3	451	6	ABU0447		Abu0447 Human exp
10	651	26.3	144	2	AAY02348		Aay02348 Protein D
11	606	24.5	446	2	AAY06590		Aay06590 Lipoprote
12	606	24.5	324	2	AAY02352		Aay02352 A represe
13	606	24.5	324	4	AAG63234		Aag63234 Amino aci
14	606	24.5	324	7	ADB67693		Adb67693 HIV-1 Lip
15	597	24.1	326	2	AAY02354		Aay02354 A represe
16	597	24.1	326	4	AAG63236		Aag63236 Amino aci
17	594.5	24.0	326	7	ADB67697		Adb67697 HIV-1 Pro
18	594.5	24.0	411	4	AAG63235		Aag63235 Amino aci
19	594.5	24.0	411	7	ADB67695		Adb67695 HIV-1 Lip
20	588.5	23.8	411	2	AAY02353		Aay02353 A represe
21	585.5	23.6	413	2	AAY02355		Aay02355 A represe
22	585.5	23.6	413	4	AAG63237		Aag63237 Amino aci
23	585.5	23.6	413	7	ADB67699		Adb67699 HIV-1 Pro
24	565	22.8	220	2	AAY25375		Aay25375 HPV fusio
25	565	22.8	220	2	AAY25378		Aay25378 HPV fusio

ALIGNMENTS

RESULT 1
 AAY79993

ID AAY79993 standard; protein; 464 AA.

XX AAY79993;

AC DT 15-MAY-2000 (first entry)

XX XX Plasmid LPB-LBI-III protein sequence.

DB XX Vaccine; non-typeable Haemophilus influenzae; nHi; infection;

KW chimeric protein; Haemophilus influenzae; P5-like fimbrial protein;

KW conjunctivitis; immunogen; antigenic; otitis media; sinusitis; conjunctivitis; lower respiratory tract infection.

XX Haemophilus influenzae.

OS Synthetic.

OS PN WO9964067-A2.

XX XX

PD 16-DEC-1999.

XX XX 99WO-US011980.

PF 28-MAY-1999;

PR 11-JUN-1998;

XX 9BGB-00012613.

(SMIK) SMITHLINE BECHMAN BIOLOGICALS.
 (OHIS) UNIV OHIO STATE RES FOUND.

PA XX Bakaletz LO, Cohen J, Dequesne G, Lobet Y;

PI XX

DR WPI ; 2000-116457/10.

DR N-PSDB; AAZ91252.

XX Novel antigenic P5-like fimbrial subunit peptides used in vaccines against Haemophilus influenzae.

XX Claim 14: Fig 5; 68pp; English.

CC The present invention describes antigenic P5-like fimbrial subunit peptides (LBI (E) peptides) of P5-like fimbrial proteins from various Haemophilus influenzae strains. The peptides are used for diagnosis, prevention, and treatment of Haemophilus influenzae infections, such as otitis media, sinusitis, conjunctivitis, or lower respiratory tract infection. The peptides may also be used in vaccines against H. influenzae. Antibodies and probes from the present invention can be used for diagnosis of H. influenzae infection. AAY7995 to AAY7993, and AAZ91201 to AAZ91252, represent sequences used in the exemplification of

AAZ91201 to AAZ91252, represent sequences used in the exemplification of

CC the present invention
 XX Sequence 464 AA;

Query Match 100.0%; Score 2477; DB 3; Length 464;
 Best Local Similarity 100.0%; Pred. No. 5.1e-207;
 Matches 464; Conservative 0; Mismatches 0; Gaps 0;

Qy 1 MDPKTKLALISLLAAGVLAGCSSHSSMANANTQMSDKLIIIAIRGASGYLPERTLESKALAPA 60
 Db 1 MDPKTKLALISLLAAGVLAGCSSHSSMANANTQMSDKLIIIAIRGASGYLPERTLESKALAPA 60
 Qy 61 QOADYLEQDLANTIKGRUVIHDHEDLGDTVAKEPFRHKRDGYVVIDPTLKETOSLE 120
 Db 61 QOADYLEQDLANTIKGRUVIHDHEDLGDTVAKEPFRHKRDGYVVIDPTLKETOSLE 120
 Qy 121 MTENFETKDGKAQAOQVYNPRFLWPKSHFRINTFEDIEFIOGLEKSTGKVGYIPEIKAPW 180
 Db 121 MTENFETKDGKAQOVPRFLWPKSHFRINTFEDIEFIOGLEKSTGKVGYIPEIKAPW 180
 Qy 181 FHHQNGKDIAAEFLKVKYGYDKCKDMVYLQTFDNEFLKRKTEFLPKGMMDLKLVQLI 240
 Db 181 FHHQNGKDIAAEFLKVKYGYDKCKDMVYLQTFDNEFLKRKTEFLPKGMMDLKLVQLI 240
 Qy 241 AYTDWKETQEKDPKGTWVNINNDWMEKPGAMEEVVYKADGVPGWTMVLVKEESKPDNIV 300
 Db 241 AYTDWKETQEKDPKGTWVNINNDWMEKPGAMEEVVYKADGVPGWTMVLVKEESKPDNIV 300
 Qy 301 YTPLVKELAQINVEVHPYTVRKDALPEFTDVNQMYDALLNKSGATGVFTDFPDGVEFL 360
 Db 301 YTPLVKELAQINVEVHPYTVRKDALPEFTDVNQMYDALLNKSGATGVFTDFPDGVEFL 360
 Qy 361 KGIKSMQGKAGVALVSQDYLYNKNISSNSTLKGHEHRARAMDGKAGCVALVRSYK 420
 Db 361 KGIKSMQGKAGVALVSQDYLYNKNISSNSTLKGHEHRARAMDGKAGCVALVRSYK 420
 Qy - 421 FYEDANGTRDHKKGRHTARTSRSDFKPYDNKRDTSFGHHHH 464
 Db 421 FYEDANGTRDHKKGRHTARTSRSDFKPYDNKRDTSFGHHHH 464

RESULT 2
 AAR20108
 ID AAR20108 standard; protein; 364 AA.
 AC AAR20108;
 DT 01-APR-1992 (first entry)
 DB Protein D - human IgD-receptor.
 XX Immunoglobulin; binding protein; probe; detection; vaccine.
 XX Haemophilus influenzae.

XX Key Location/Qualifiers
 PH
 FT Peptide
 PT Peptide
 PN WO9118926-A.
 PD 12-DEC-1991.
 XX 31-MAY-1990; 90SE-00001949.
 XX 31-MAY-1990; 90SE-00001949.
 XX (FORS/) FORSGREN A.
 XX FORsgren A;

OS Haemophilus influenzae sequence"
 XX /note= "bacterial lipoprotein consensus sequence"

RESULT 3
 DR WPI; 1992-007432/01.
 DR N-PSDB; AAQ20265.
 PT Surface-exposed protein conserved in strains of *Haemophilus influenzae* -
 PT useful in vaccines, and antibodies to protein, are used to detect
 PT Haemophilus in samples.
 PS Disclosure: Fig 9a-b; 42pp; English.

XX The surface-exposed protein, conserved in many strains of *H. influenzae* or related *Haemophilus* species, has a mol. wt. of 42 kD and a binding capacity for human IgD. The DNA sequence can be used to design probes and primers for the detection of *H. influenza* and related species. The protein can be used to detect and isolate IgD and to prepare vaccines against *H. influenzae*
 XX Sequence 364 AA;
 Query Match 77.5%; Score 1920; DB 2; Length 364;
 Best Local Similarity 99.5%; Pred. No. 1.4e-158;
 Matches 362; Conservative 0; Mi smatches 2; Indels 0; Gaps 0;
 Qy 1 MDPKTKLALISLLAAGVLAGCSSHSSMANANTQMSDKLIIIAIRGASGYLPERTLESKALAPA 60
 Db 1 MKLKTIALSLLAAGVLAGCSSHSSMANANTQMSDKLIIIAIRGASGYLPERTLESKALAPA 60
 Qy 61 QOADYLEQDLANTIKGRUVIHDHEDLGDTVAKEPFRHKRDGYVVIDPTLKETOSLE 120
 Db 61 QOADYLEQDLANTIKGRUVIHDHEDLGDTVAKEPFRHKRDGYVVIDPTLKETOSLE 120
 Qy 121 MTENFETKDGKAQAOQVYNPRFLWPKSHFRINTFEDIEFIOGLEKSTGKVGYIPEIKAPW 180
 Db 121 MTENFETKDGKAQOVPRFLWPKSHFRINTFEDIEFIOGLEKSTGKVGYIPEIKAPW 180
 Qy 181 FHHQNGKDIAAEFLKVKYGYDKCKDMVYLQTFDNEFLKRKTEFLPKGMMDLKLVQLI 240
 Db 181 FHHQNGKDIAAEFLKVKYGYDKCKDMVYLQTFDNEFLKRKTEFLPKGMMDLKLVQLI 240
 Qy 241 AYTDWKETQEKDPKGTWVNINNDWMEKPGAMEEVVYKADGVPGWTMVLVKEESKPDNIV 300
 Db 241 AYTDWKETQEKDPKGTWVNINNDWMEKPGAMEEVVYKADGVPGWTMVLVKEESKPDNIV 300
 Qy 301 YTPLVKELAQINVEVHPYTVRKDALPEFTDVNQMYDALLNKSGATGVFTDFPDGVEFL 360
 Db 301 YTPLVKELAQINVEVHPYTVRKDALPEFTDVNQMYDALLNKSGATGVFTDFPDGVEFL 360
 Qy 361 KGIKSMQGKAGVALVSQDYLYNKNISSNSTLKGHEHRARAMDGKAGCVALVRSYK 420
 Db 361 KGIKSMQGKAGVALVSQDYLYNKNISSNSTLKGHEHRARAMDGKAGCVALVRSYK 420
 Qy - 421 FYEDANGTRDHKKGRHTARTSRSDFKPYDNKRDTSFGHHHH 464
 Db 421 FYEDANGTRDHKKGRHTARTSRSDFKPYDNKRDTSFGHHHH 464

DR DR 24-FEB-2003 (first entry)
 DR ABB84614 standard; protein; 358 AA.
 XX AC ABB84614;
 XX DT 24-FEB-2003 (first entry)
 XX E. coli glycerophosphoryl phosphodiester phosphodiesterase.

XX KW antihaemic; cerebroprotective; immunosuppressive; dermatological; cardiotonic; antinflammatory; cytostatic; immunostimulant; antit-HIV; neuroprotective; KW tranquilliser; nephrotoxic; thymomimetic; gene mapping; apoptosis; AIDS; cell cycle disruption; programmed cell death regulation; KW viral infection; nucleosome assembly; phosphatase homeostasis; cell cycle regulation; cancer; follicular lymphoma; carcinoma; KW p53 mutation; graft rejection; hormone-dependent tumour; autoimmune disorder; valvular heart disease; systemic lupus erythematosus; diabetes; Hashimoto's thyroiditis; virus-induced lymphocyte depletion; KW immune-mediated glomerulonephritis; KW

acquired immunodeficiency syndrome; neurodegenerative disease; stroke;
 KW Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
 KW spinal muscular atrophy; retinitis pigmentosa; cerebellar degeneration;
 KW myelodysplastic syndrome; ischaemic injury; myocardial infarction;
 KW reperfusion injury; liver disease; idiopathic dilated cardiomyopathy;
 KW ischaemic cardiomyopathy; aplastic anaemia; chronic neutropenia; mania;
 KW myelodysplastic syndrome; central nervous system disorder; anxiety;
 KW senile dementia; Huntington's disease; hypertension; schizophrenia;
 KW severe bipolar affective disorder;

KW glycerophosphoryl phosphodiester phosphodiesterase.

XX OS Escherichia coli.

PN WO200281516-A2.

XX PD 17-OCT-2002.

XX PF 16-JAN-2002; 2002WO-US001098.

XX PR 16-JAN-2001; 2001US-0262206P.

PR 15-JAN-2002; 2002US-00047855.

XX PA (MILLI-) MILLENNIUM PHARM INC.

PI Chiang LW;

XX WPI: 2003-058503/05.

XX Novel isolated programmed cell death-related polypeptide, NARC10 and
 NARC16, useful for treating disorders associated with abnormal apoptotic
 process e.g. Alzheimer's disease, cancer, myocardial infarction, stroke.

PS Disclosure; Fig 3; 123pp; English.

XX This invention describes novel cell death-related polypeptides NARC10 and
 NARC16, located on chromosome 4q11-q21 and which have cardiotropic, antihIV,
 immunosuppressive, dermatological, antiinflammatory, cerebroprotective,
 CC neuroprotective, antianemic, cardiotropic, vasoconstrictive, antihemolytic,
 CC antidiabetic, immunosuppressive, cyclostatic, thymomimetic, nephrotoxic,
 CC immunostimulant, anticonvulsant, tranquiliser, hypotensive and
 CC neuroleptic activity and can be used in gene therapy. The products of the
 CC invention can be used to modulate NARC10 or NARC16 polypeptides or
 CC polynucleotides, to map NARC genes on a chromosome, e.g. to locate gene
 CC regions associated with genetic disease or to associate NARC10 or NARC16
 CC with a disease. The polypeptides are also useful for modulating the
 CC apoptotic process, and are therefore useful for modulating, and treating
 CC disorders associated with increased apoptosis, inhibition of apoptosis or
 CC disruptions in cell cycle, for regulating cellular functions including
 CC programmed cell death, nucleosome assembly, phosphate homeostasis and the
 CC cell cycle. Preferably, the products of the invention are useful for
 CC treating disorders associated with abnormally low rate or abnormally high
 CC rate of apoptosis e.g. cancers including follicular lymphomas, carcinomas
 CC with p53 mutations, or hormone-dependent tumours, autoimmune disorders
 CC including systemic lupus erythematosus, diabetes, graft rejection,
 CC Hashimoto's thyroiditis and immune-mediated glomerulonephritis and viral
 CC infections e.g. infections caused by herpes viruses, virus-induced
 CC lymphocyte depletion (including acquired immunodeficiency syndrome
 CC (AIDS)), neurodegenerative diseases manifested by loss of specific sets
 CC of neurons (including Alzheimer's disease, Parkinson's disease,
 CC amyotrophic lateral sclerosis, spinal muscular atrophy, retinitis
 CC pigmentosa, and cerebellar degeneration), myelodysplastic syndromes
 CC (including aplastic anaemia), ischaemic injuries (including myocardial
 CC infarction, stroke and reperfusion injury), and toxin (e.g. alcohol)
 CC induced liver disease, idiopathic dilated cardiomyopathy, ischaemic
 CC cardiomyopathy and valvular heart disease, aplastic anaemia, chronic
 CC neutropenia, and myelodysplastic syndromes, central nervous system
 CC disorders, senile dementia, Huntington's disease, hypertension,
 CC schizophrenia, attention deficit disorder, mania, anxiety, severe bipolar
 CC affective disorder (BP-1). This sequence represents the E. coli
 CC glycerophosphoryl phosphodiester phosphodiesterase protein described in
 CC the method of the invention.

XX Sequence 358 AA:

Query Match	50 5%	Score 1251.5;	DB 6;	Length 358;
Best Local Similarity	64.1%	Pred. No. 2.7e-100;		
Matches 229;	Conservative 56;	Mismatches 65;	Indels 7;	Gaps 2;
Qy	4 KTLASLLANGVLAGSSHSNNMANNTQMSDKIIIAHGRASGYLDEHTLESKALPAQQA	63		
Db	6 KNLSNAIMSTIVMSANAAAD-----SNEKIVIAHGRASGYLFPEHTLPAKAMAYAQQA	59		
Qy	64 DYLEBDLAMPTKDGRVIVTHPHFDLDTDYAKKPFRHRKGRTYIDFTKEIQSIQEMTE	123		
Db	60 DYLBQDLMVTKDDNLYLVLDHYLDWTVDARFPRBARDGRTYIDFILDEIKSLKFT	119		
Qy	124 NFETTDKGKOAQVYPFRPLWKSHPRINTPDEIEPIQGLEKSTGKVGTYPEIKAPWFHH	183		
Db	120 GFDINGRKVQTYPRFPMSDFPVHTPFBIEEVQGLNHSTGNRNGTYPEIKAPWFHH	179		
Qy	184 QNGKQIAAETFLKULKKYGDKKTMVYLQTPDFNLKRIKTELLQMGMDKLVLQVIAVT	243		
Db	180 QBGRQDAAKTLLEVKKYGTGKODKRYLQCFDADLKRNELBPRGMHLNVQLVIAVT	239		
Qy	244 DWKETQEKDQDGRVYWWNNYNDWMFKRGAIAEVVKYTAQGVGIGWMMVUNKBEISKPDIVYTP	303		
Db	240 DWNETQQKQDPGSWNNYNDWMFKRGAQVAEFDGIGBDYHMLI-BETSQPGNKLGTG	298		
Qy	304 LVKELLAQYNVEVHVPYTVRKDALPEPTDYNQMDLLNSKGATGVFTDFTDGTGVRPL	360		
Db	299 MVQDAQQNKLVWHPYTVRSDDLPETTPDVNQLYDLYNKAGVNGI-LFTDFDKAVRPL	355		

RESULT 4

ID ABO61848 standard; protein; 359 AA.

XX XX

XX AC ABO61848;

XX DT 29-JUL-2004 (first entry)

XX DB Klebsiella pneumoniae polypeptide seqid 8365..

XX Recombinant expression vector; transcription regulatory element;
 KK Klebsiella pneumoniae protein; antibacterial; Vaccine.

XX OS Klebsiella pneumoniae.

XX PN US6610836-B1.

XX PD 26-AUG-2003.

XX PP 27-JAN-2000; 2000US-00489039.

XX PR 29-JAN-1999; 99US-011747P.

XX PA (GENO-) GENOME THERAPEUTICS CORP.

XX PI Bretton GL, Osborne M;

XX DR WPI: 2003-895146/82.

XX DR N-PSDB; ACH95399.

XX PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
 XX preparing a vaccine composition against Klebsiella pneumoniae.

XX PS Disclosure; SEQ ID NO 8365; 932pp; English.

XX CC The invention describes a new isolated nucleic acid encoding a Klebsiella
 CC pneumoniae polypeptide. Also described are: a recombinant expression
 CC vector comprising the nucleic acid, operably linked to a transcription
 CC regulatory element, and a cell comprising the recombinant expression
 CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against Klebsiella pneumoniae. This is the amino acid sequence of a
 CC Klebsiella pneumoniae polypeptide of the invention

XX XX

SQ	Sequence 359 AA;	CC polypeptide and a method for screening test compounds for anti-bacterial activity. The polypeptides and polynucleotides are useful as molecular targets for diagnosing, preventing and treating pathological conditions resulting from bacterial infection, as reagents for diagnosis of bacterial diseases, as components of antibacterial vaccines, or as targets for antibacterial drugs or as bio-control agents for plants. This sequence represents a Proteus mirabilis polypeptide of the invention.
Query Match	50.5%; Score 1251.5; DB 7; Length 359;	CC
Best Local Similarity	65.3%; Pred. No. 2.7e-100;	CC
Matches	231; Conservative 48; Mismatches 74; Indels 1; Gaps 1;	CC
Qy	7 ALSLLAAGVLLAGCSSHSSNMANANTOMKSDFKIIIAHRAASGYLPEFTLESKALAFAAQADYL 66	CC
: Db	8 AMKMTKTAIAMSOMISSLSSAECATADKMYIAHRAASGYLPEFTLESKALAFAAQADYL 62	CC
Query	67 EQDLAMTKDGRLLVYTHIDHEFLGLTDAVAKFPHRKDGRRYVIDFTKETQSLENTENPE 126	CC
: Db	63 EQDLWTRDRLVBLHYDRLVDTDAQRFFORAKRDKGRFAIDFTLDEIKSLKETEGE 122	CC
Query	127 TKDGKQAOQYENPNEPFLMKSHPRHTFEDTEFLQGLEKSTGKVKGYYPEKAPNPFHQNG 186	CC
: Db	123 PRNGKNVQTYPGRFPWCKSDPRHTEEIRFVQGLNHSTCKNIGTYPEKAPNPFHQSG 182	CC
Query	187 KDIAAETLKVLLKKGYDKTKDMVYLQTFDFNELLKRITKEPLLQMGNDLKVQLIAYTDWK 246	CC
: Db	183 KDIAAETLKVLLKKGYDKTKDMVYLQTFDFNELLKRITKEPLLQMGNDLKVQLIAYTDWN 242	CC
Query	247 ETQEKDPKGWYNNYNYDMWMPFKPGAMAEVVKYADGVGPWYMLVNKEESKEDPNIVYTPLYK 306	CC
: Db	243 ETQKQADGKRWNNYSDMWFKEGAMAQIAQADGQDIAHLV-AEGSKPGAVKLTMVK 301	CC
Query	307 ELAQXNWVHPTVTKDALLPEFTDQNMDALLNSGATGSVTFEPDTSYFEL 360	CC
: Db	302 EAHASHLQVHPYTVRDAQLPQEPATANVNLQDLYNQAGVDSLFTFPDKAVQFL 355	CC
RESULT 5	248 TQEKPDKGWTWNTNNYDMWMPKGAMAEVVKYADGVGPWYMLVNKEESKEDPNIVYTPLYK 307	CC
ID ADFO5718	248 TYRSPDPGTWNTNNYDMWMPKGAMKETIATADGIPDQYHMLV-BEDSTDPKTLTGMAAD 306	CC
XX	248 ABM68758 standard; protein; 359 AA.	CC
AC ADFO5718;	248 ABM68758;	CC
XX	20-NOV-2003 (first entry)	CC
DT 12-FEB-2004 (first entry)	20-NOV-2003 (first entry)	CC
DB Bacterial polypeptide #1831.	DB Photorhabdus luminescens protein sequence #1855.	CC
KW Proteus mirabilis infection; bacterial infection; antibiotic; immunostimulant.	KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis; detection; food; gene expression; plant; animal; microorganism; toxin; ant biotic; biopesticide; virulence factor; disease model; plague; whooping cough.	CC
KW	KW	CC
OS Proteus mirabilis.	OS Photorhabdus luminescens.	CC
XX	XX	CC
PN US6605709-B1.	PN WO200294867-A2.	CC
XX	XX	CC
PD 12-AUG-2003.	PD 28-MOV-2002.	CC
XX	XX	CC
PP 05-APR-2000; 2000US-00543681.	PP 07-FEB-2002; 2002WO-IB003040.	CC
XX	XX	CC
PR 09-APR-1999; 99US-0128706P.	PR 07-FEB-2001; 2001FR-00001659.	CC
XX	XX	CC
PA (GENO-) GENOME THERAPEUTICS CORP.	PA (INSP) INST PASTEUR.	CC
XX	PA (CNRS) CNRS CENT NAT RECH SCI.	CC
PI Bretton GL;	XX	CC
XX	XX	CC
WPI; 2003-895291/82.	XX	CC
DR N-PSDB; AD01546.	XX	CC
XX	XX	CC
New Proteus mirabilis polypeptides and polynucleotides, useful as reagents for diagnosis of bacterial disease, as components of antibacterial vaccines, as targets for antibacterial drugs, or as biocontrol agents for plants.	Antibacterial; fungicide; insecticide; polymorphism; genetic analysis; detection; food; gene expression; plant; animal; microorganism; toxin; ant biotic; biopesticide; virulence factor; disease model; plague; whooping cough.	CC
Disclosure; SEQ ID NO 6003; 870pp; English.	CC	CC
XX	XX	CC
The invention relates to new Proteus mirabilis polypeptides and polynucleotides. The invention also relates to antibodies against the polypeptides. The methods for producing the polypeptides, a method of generating vaccines for immunising an individual against P. mirabilis, a method for evaluating a compound for the ability to bind a P. mirabilis	The invention relates to new Proteus mirabilis polypeptides and polynucleotides. The invention also relates to antibodies against the polypeptides. The methods for producing the polypeptides, a method of generating vaccines for immunising an individual against P. mirabilis, a method for evaluating a compound for the ability to bind a P. mirabilis	CC
CC	CC	CC
PT	PT	CC
PT	PT	CC
XX	XX	CC
PS	Duchaud B, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A; Buchrieser C;	CC
XX	XX	CC
DR	WPI; 2003-148459/14.	CC
XX	XX	CC
PT Genomic sequence of Photorhabdus luminescens and encoded polypeptides.	Genomic sequence of Photorhabdus luminescens and encoded polypeptides.	CC

useful e.g. as therapeutic antimicrobials and agricultural pesticides.

XX Claim 2; SEQ ID NO 1855; 1205pp; French.

CC The invention relates to the isolation of genes and their encoded proteins from *Photinus* *luminosus*. The isolated sequences are sources of probes and primers for detecting the genome of *P. luminosus* and related species; to study polymorphisms; for gene analysis and for detection/amplification of the genes. Antibodies (Ab) raised against the polypeptides encoded by the genes are used for detection/identification of *P. luminosus*, e.g. in foods. The genes (proteins, Ab and cells that carry a gene-containing vector are used to select compounds that modulate, regulate, induce or inhibit expression of the genes in plants, animals or microorganisms other than *P. luminosus* and are able to alter response or sensitivity to toxins and antibiotics produced by *P. luminosus*. Cells transformed to express the genes are useful for recombinant production of the proteins, particularly toxins and antibiotics useful as insecticides, bactericides and fungicides. The genes, proteins, vectors containing the genes and Ab are also useful therapeutically (to treat microbial infection by bacteria or fungi that are sensitive to *P. luminosus*-encoded toxins or antibiotics) and as bioprecides. Other uses of the genes and the proteins are as virulence factors and for identifying targets of human diseases for which *P. luminosus* is a model (particularly plague and whooping cough). This sequence represents one of the isolated *P. luminosus* proteins

Sequence 359 AA;

Query Match 48.9%; Score 1210.5; DB 6; Length 359;

Best Local Similarity 62.0%; Pred. No. 1e-96; Matches 54; Mismatches 71; Gaps 11; Indels 11; Gaps 2;

Matches 222; Conservative 54; Substitutions 190; Insertions 1; Deletions 1; Gaps 1;

Query 4 KTIALSLLAAAGVLAGCSSHSMANTOMKSIIIAHGRASCYLPPEHTLESKALAPAQAA 63

Db 6 KTMIGIGIL-----TSSMSGTAQAAKIVIAHGRASCYLPPEHTLPKAMAYEGA 55

Query 64 DYLEQDLAMTKDGRLLVYTHDHPGLTDVACKCPHRHRKGDRYYVIDFTLKEIQSLMTE 123

Db 56 DYLEQDLWTKDDELIVHDYLDRTDVANKFPNARQDGRRYAYADFTLSEISLKFTE 115

Query 124 NFETKDQKQAQVYQPNRFPWKSHERTHITFEDBIEFIQGLEKSTGKVGTYPIBKAPWFHH 183

Db 116 GFDIKNDRQIQNSNRFPWKSDFRHTFQERIEFYQGLNSTGKNTGTYPIBKAPWFHQ 175

Query 184 QNGKDIATELTKLVKKYGDKTKDMVLYLQTDPENELKRITTELLPONGMDKLKVQIAYT 243

Db 176 KEKGDISTKVLKAVGTYKSDKIVLQCFPTNELLKRNNELEPKLGMDKLKVQIAYT 235

Query 244 DMKETQEKDPKGKSYWVNTYDNMFKPGMAMEVYKYADCGPGWYMLNEKEESKRPDNIVYTP 303

Db 236 DNNETYEXQSDGKWTNYSYDNMFKPGMAMEVYQADQGDPYHMLVEK-YSTPTNIKLN 294

Query 304 LYKELAQNYEVHYPYTYRKDALPEFFTDVNOMYDALINKSGATGVFTDFPDGTGVFLK 361

Db 295 LYKEAHTNNLEVHYPTYRVDQLPKYATSGQDFLIIYNQAGYDGVFTDFPDLGKFLQ 352

/ RESULT 7
AAY52773 standard; protein; 356 AA.

AC AAY52773;
XX AAY52773;

XX 26-JAN-2000 (first entry)

DE Treponema pallidum glycerophosphodiester phosphodiesterase.

KW Treponema pallidum; syphilis; vaccine; treponemal disease; Msp; Yaws; major sheath protein; bejai; gingivitis; periodontal disease; pinta.

OS Treponema pallidum.

PN WO9953099-A1.

PT	XX	PD 21-OCT-1999.
PS	XX	PF 09-APR-1999; 99WO-US007886.
CC	XX	PR 10-APR-1998; 98US-00058968.
CC	XX	(UNIW) UNIV WASHINGTON.
CC	XX	PA Van Voorhis WC, Lukehart SA, Centurion-Lara GA, Cameron CES;
CC	XX	PI WPI; 1999-62045/53.
CC	XX	DR N-PSDB; AAZ33104.
CC	XX	PT Novel proteins useful in vaccines against syphilis and other treponemal diseases.
CC	XX	PS Claim 1; Page 94-95; 200pp; English.
CC	CC	The present invention describes novel <i>Treponema pallidum</i> genes. These genes encode a glycerophosphodiester phosphodiesterase (Gpd), a D15/Cma87 homologue, and proteins with homology to major outer sheath (Msp) proteins of <i>T. denticola</i> . Also described are: (1) an isolated protein capable of inducing a protective immunologic response to <i>T. pallidum</i> , <i>T. p. pertenue</i> , or <i>T. p. endemicum</i> , when administered in an effective amount to an animal host; (2) a method of identifying a <i>T. p. pallidum</i> vaccine candidate; (3) a method of inducing a protective immune response against <i>T. pallidum</i> ; (4) a method for analysing a sample of DNA to determine whether it originated from <i>T. p.</i> subspecies <i>pallidum</i> , <i>T. p.</i> subspecies <i>pertenue</i> or <i>T. p.</i> subspecies <i>endemicum</i> , and (5) a method of determining whether a first and a second clinical isolate of <i>T. p.</i> <i>pallidum</i> are the same or different. The proteins are used, either alone or in combination, in vaccines against Treponemal diseases, e.g. syphilis, bejai, pinta, yaws, gingivitis, and periodontal disease. They may also provide protection against other Treponemal diseases. The methods may be used to identify vaccine candidates, and to determine the origin of a treponemal nucleic acid. AAY52773 to AAZ33104 to AAZ33156, and AAY52773 to AAY52811, represent sequences used in the exemplification of the present invention.
CC	CC	SQ Sequence 356 AA;
CC	CC	Query Match 43.4%; Score 1074; DB 2; Length 356;
CC	CC	Best Local Similarity 54.6%; Pred. No. 8.1e-85; Mismatches 67; Missmatches 79; Indels 12; Gaps 1;
CC	CC	Matches 190; Conservative 67; AALVACCA-----SERMVAYRGRAGTYPEHTPAQGDAYDLYQDGVVL 62
CC	CC	Query 13 AGVLAGCSSHSSNMANQMSDKI1IAHGRASGYLPEHTLESKALAFQADQYLEDQDLAM 72
CC	CC	Db 15 TKDGHLYVHDHFGLCLTDYAKKPFRHRKGDRGTVYDIFPLIKEIQLSLENTEFETKDGKQ 132
CC	CC	Query 63 SKDNQOLIVAQSIIHLNMDTDAEKFPRQRORADGHPFTVIDFTVEELSLRATNTSFYTRGKRH 122
CC	CC	Db 63 AQVYVNPFRPLWKSFRHTPEDEIPTGQLEKSTGKVGTYPIBKAPWFHHQNGDIAAE 192
CC	CC	Query 133 TPIVIGORFRPLWKSFRHTPEDEIPTGQLEKSTGKVGTYPIBKAPWFHHQEGDIAAL 182
CC	CC	Db 123 193 TLKVLKKYGYDQKTKDQVYDQFDENLKRITKELLPQMGMDLKLVQIAYT 252
CC	CC	Query 193 SRGKVNNTNNWMMFPEFGGMQKIAKJADGCVPGWYMLVNKEESKPDNNIVTPVKELAQYN 312
CC	CC	Db 183 TLALRKYGYQSRSLVYVOTYDFBLKRITKELLPKYENKVLJQRVATDQRSQEQD 242
CC	CC	Query 253 PKGWWNTNNWMMFPEFGGMQDPMVSIQDQK 302
CC	CC	Db 243 VEVHPVTVRDALPEFFTVDQMYDALLNKSGATGVFTDPDTGVBF 360
CC	CC	Query 313 DE Treponema pallidum glycerophosphodiester phosphodiesterase.
CC	CC	Db 303 LECHVATVRKTELPSYARTMDEMFSTLKFQTGAVNVLTDFDGLKFL 350

RESULT 8
AAY06589

ID	AAY06589	standard; protein; 451 AA.	Qy	121	MTEFETDKQKAOQVYPNRFPLWKSHFRINTFEDIEFQGLEKSFGKRVGIVYPEIKAPW 180
XX			Db	121	MTEFETDKQKAOQVYPNRFPLWKSHFRINTFEDIEFQGLEKSFGKRVGIVYPEIKAPW 159
AC					
XX	26-OCT-1999	(First entry)	Qy	181	FHHONGDIAETTKVLKRYGYDKRTDMYLOTFDNEFKRTEPQGMDFKLVL 239
DT			Db	160	TEEQEAAASSSTLVEV-----TLEGEPAAESPDPPSPQGASSLPT 201
XX	Lipoprotein D-MAGE-3-His fusion protein.				
DE					
XX	MAGE-3; lipoprotein D; LPD-MAGE-3-His; fusion protein; tumour; melanoma; breast cancer; bladder cancer; lung cancer; head and squamous cell carcinoma; colon cancer; oesophagus carcinoma; vaccine; human.		Qy	240	IAYTWKETOB---KDPKGWYNNYNYDPMFKPG---AMAEVV---AD 279
FT			Db	202	MNPFLWSQS TEDSSNQEEBGPSPTPDLSBFQAAISRKVAELVHFLKYRAREPVTKA 261
FT	Haemophilus influenzae.		Qy	280	GVGP--GWYMLVNKEESKPDNIVYTPLVKEAQNVETHAPYTRKDALPEPTFDVNQTY 336
OS	Homo sapiens.		Db	262	MIGSVGMMQYFFPVIFPSASSQQLVFCIELMEDPFLGHLYI-----PATCGLSY 313
OS	Synthetic.		Qy	337	DALLMKGATGVFTFPDTGVEFL-KGIKSMDGRAGVALVRSDFLY----NRNSSN 390
OS	Chimeric.		Db	314	DGLLGDNQI----MPKAGLLIVTLAIREGDCAPEEKIWEELSYLEVPEGREDLSILG 367
XX			Qy	391	STUKNLGEHHRARAMGGKAGVALVRSDFKYEDANG-----TRDHCK 433
PH	Key Misc-difference 43	Location/Qualifiers	Db	368	DPKCLLTOH-----FVQENTLEYROPGSDPACYBFLWGDRALVETSYKV 413
FT	/note= "this residue is additional to the residues deduced from the nucleotide sequence of AAX87588"		Qy	434	GRIHTARTSRSDYKFD-----NKRDTSGHHHHHH 464
FT			Db	414	LHRMKVKGCGPHISYPPPLHEWLREGEEFTGGHHHHHHH 451
PN	W09940188-A2.				
XX					
PD	12-AUG-1999.				
XX					
PF	02-FEB-1999;	99WO-EP000660.			
XX					
PR	05-FEB-1998;	98GB-00002543.			
PR	06-FEB-1998;	98GB-00002650.			
XX					
PA	(SMIK) SMITHKLINE BEECHAM BIOLOGICALS.				
XX					
PI	Cabezon Silvia T,	Cohen J,	Qy	RESULT 9	
XX			ABU04447	ABU04447 standard; protein; 451 AA.	
DR	WPI; 1999-494293/41.		ID	ABU04447	
DR	N-PSDB; AAX87588.		DT	29-JAN-2003 (first entry)	
XX			XX		
XX	New protein derivatives used in cancer vaccine therapy for treating a range of cancers including melanomas, carcinomas and cancers of breast.		XX		
PT	Example 1; Page 64-65; 72pp; English.		XX		
XX	This sequence represents a novel fusion protein composed of lipidated protein D (LPD) of Haemophilus influenzae B, the human MAGE-3 tumour-associated antigen and a hexahistidine tail. A vector designed for recombinant expression of the fusion protein is provided. MAGE-3 cDNA was amplified using primers that altered the first 5 codons to Escherichia coli codon usage. The LPD moiety provided the fusion protein with additional exogenous T-cell epitopes and also increased expression levels in E. coli. The lipid tail ensured optimal presentation of the antigen to antigen-presenting cells. The affinity tag fused to an immunological fusion partner, e.g. LPD-MAGE-3-His. These novel fusion proteins provide vaccines for immunotherapy of melanomas or other MAGE-associated tumours like breast, bladder, lung and non-small cell lung cancer, head and squamous cell carcinoma, colon carcinoma and oesophagus carcinoma				
CC	Sequence 451 AA;		XX		
CC	Query Match 26.7%; Score 661.5; DB 2; Length 451;		XX		
CC	Best Local Similarity 37.5%; Pred. No. 1.1e-48;		XX		
CC	Matches 194; Conservative 51; Mismatches 152; Indels 121; Gaps 18;		XX		
CC	WPI; 2003-040507/03.		DR		
CC	1 MDPKTLAISLILLAGVLAGCSSHSNNANTQMSDKIIIAHRGASGYLPEHTLESKALAPA 60		XX		
CC	1 MDPKTLAISLILLAGVLAGCSSHSNNANTQMSDKIIIAHRGASGYLPEHTLESKALAPA 60		PT	New polypeptides (e.g. kinases, phosphatases, proteases, transporters, receptors or transcription factors), useful for	
CC	QADYLLEQDLAMTIKDGRLVVTHDHELDLGTVAKKEPHHRKDGRRYYVIFTLKIQSQLE 120		PR	Cytoskeletal Proteins, 1-MAY-2001; 2001US-0292544P.	
CC	QADYLLEQDLAMTIKDGRLVVTHDHELDLGTVAKKEPHHRKDGRRYYVIFTLKIQSQLE 120		PR	08-AUG-2001; 2001US-0310801P.	
CC	QADYLLEQDLAMTIKDGRLVVTHDHELDLGTVAKKEPHHRKDGRRYYVIFTLKIQSQLE 120		PR	01-OCT-2001; 2001US-0326370P.	
CC	QADYLLEQDLAMTIKDGRLVVTHDHELDLGTVAKKEPHHRKDGRRYYVIFTLKIQSQLE 120		PR	04-DEC-2001; 2001US-0336780P.	
CC	QADYLLEQDLAMTIKDGRLVVTHDHELDLGTVAKKEPHHRKDGRRYYVIFTLKIQSQLE 120		PR	20-FEB-2002; 2002US-0355985P.	
SQ	Sequence 451 AA;		XX		
SQ	ZYCO - ZYCOS INC.		XX		
SQ	Chicz RM, Tonlinson AJ, Urban RG;		XX		
Db			XX		
Qy			XX		
Db			XX		
Qy			XX		
Db			XX		

The invention describes a purified polypeptide, which comprises a fragment of a kinase, phosphatase, protease, protease inhibitor, transporter, cytoskeletal protein, receptor or transcription factor. The polypeptide is useful as an immunogenic composition for eliciting in a mammal an immunogenic response directed against any of the purified polypeptide. The purified polypeptide, or the antibody that binds to this polypeptide, is useful for treating cancer. The polypeptide is also useful for identifying compounds that binds to a naturally processed class I or class II MHC binding polypeptide. The polypeptides and nucleotides are particularly useful for treating or preventing myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma, lymphoma or leukaemia. These are also useful for screening agents for treating the above mentioned diseases. This sequence represents an expressed protein tag (EPT) isolated from human tissue for translational profiling. Note: This sequence does not appear in the printed specification but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

Sequence 451 AA:

Query Match 26.7%; Score 661.5; DB 6; Length 451;

Best Local Similarity 37.5%; Pred. No. 1.1e-48; Matches 194; Conservative 51; Indels 121; Gaps 18;

PS Example 2; Fig 1; 66pp; English.

Qy 1 MDPKTLAISLLAAGVLAGCSSSHSNMANTQMSDKIIIAHRCASGVLPEHTLESKALAF A 60
 Db 1 MDBKTLAISLLAAGVLAGCSSSHSNMANTQMSDKIIIAHRCASGVLPEHTLESKALAF A 60

Qy 61 QQADYLBEDLANTKDGRVVTHDFLDGLTDVAKKFPHRKDGRTYVIDFTLKEIQSLE 120
 Db 61 QQADYLBEDLANTKDGRVVTHDFLDGLTDVAKKFPHRKDGRTYVIDFTLKEIQSLE 120

Qy 121 MTENFETDGDQAQVYPNRFLPKWHSRINTFDEIEFIQLEKSTGKKGKTYPEIKA PW 180
 Db 121 MTENFETDGDQAQVYPNRFLPKWHSRINTFDEIEFIQLEKSTGKKGKTYPEIKA PW 180

Qy 181 FHHQNGKDIQAETTLKVKRYGYDKKTDMVLYQTFDNELKLKTELLPQMGMDL -KLVLQL 239
 Db -

Db 160 TBEQEAAASSSSLTVEV-----AD 279
 Qy 240 IAYTDWKETOE----KDPKGYVWNVNNDWMFKPG----AMAEVV-----KY-----AD 279
 Db 202 MNPLWNSCSYEDSNQBEGPSTFDLISEFOALSRSRKAELVHFLLKTYAREPVTKAE 261

Qy 280 GWG-----GWWMLVNKEEKSXPDKNIVYPTPLKVELAQYNVNEVHPYPTVVKDALPEFTDVNQMY 336
 Db 262 MUGSVVENVWQYFPVFSKASSSLQLVFGIELMEVDPIGHLYI-----PATCIGLSY 313

Qy 337 DALLNKSGATGVTFDPPTGVBL -KGISMDGGKAGVALVRSDYK-----TRDHKK 433
 Db 314 DGLGDNQI-----MPKAGLILIVLAIARRGDCAPEKWIWBLSLVEPFEGREDISLG 367

Qy 391 STKLKLGFFHRARAMDGKAGVALVRSDYK-----TRDHKK 433
 Db 368 DPKCLLJTOH-----FVQENYLERYQVPSDPACTEFLMGPRLVETSYVKV 413

Qy 434 GRHTARTSRSDYKFYD-----NRKIDSTSCHHHHHH 464
 Db 414 LHHMVKITGGPHISYPPLHEWVIREGSETSGSHHHHHH 451

RESULT 10
 AAY02348 standard; protein; 144 AA.
 AC AAY02348;
 XX DT 17-OCT-2003 (revised)
 DT 09-JUL-1999 (first entry)

DB Protein D, the fusion partner for HIV-1 Nef/rat fusion proteins.

XX HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;

XX vaccine; HIV infection; protein D.

XX Haemophilus influenzae; B.

OS WO916884-A1.

XX 08-APR-1999.

XX 17-SEP-1998; 98WO-EPO06040.

XX 26-SEP-1997; 97GB-00020585.

XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX PI Bruck C, Godart SAG, Marchand M;

XX DR WPI; 1999-302282/25.

DR N-PSDB; AAX35654.

XX HIV Tat or Nef protein linked to a fusion partner.

XX PT PS Example 2; Fig 1; 66pp; English.

XX CC The present sequence represents protein D. This protein acts as the fusion partner for the fusion proteins of this invention, which also comprise HIV-1 Nef or Tat proteins (or derivative). The fusion protein can be used in vaccine to prevent HIV infection. (Updated on 17-OCT-2003 to standardise OS Field)

CC XX SQ Sequence 144 AA;

SQ Query Match 26.3%; Score 651; DB 2; Length 144;

SQ Best Local Similarity 100.0%; Pred. No. 1.7e-48; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDPKTLAISLLAAGVLAGCSSSHSNMANTQMSDKIIIAHRCASGVLPEHTLESKALAF A 60

Db 1 MDPKTLAISLLAAGVLAGCSSSHSNMANTQMSDKIIIAHRCASGVLPEHTLESKALAF A 60

Qy 61 QQADYLBEDLANTKDGRVVTHDFLDGLTDVAKKFPHRKDGRTYVIDFTLKEIQSLE 120

Db 61 QQADYLBEDLANTKDGRVVTHDFLDGLTDVAKKFPHRKDGRTYVIDFTLKEIQSLE 120

Qy 121 MTENFETDGDQAQVYPNRFLPKWHSRINTFDEIEFIQLEKSTGKKGKTYPEIKA PW 180

Db 121 MTENFETDGDQAQVYPNRFLPKWHSRINTFDEIEFIQLEKSTGKKGKTYPEIKA PW 180

Qy 181 FHHQNGKDIQAETTLKVKRYGYDKKTDMVLYQTFDNELKLKTELLPQMGMDL -KLVLQL 239

Db -

Qy 61 QQADYLBEDLANTKDGRVVTHDFLDGLTDVAKKFPHRKDGRTYVIDFTLKEIQSLE 120

Db 61 QQADYLBEDLANTKDGRVVTHDFLDGLTDVAKKFPHRKDGRTYVIDFTLKEIQSLE 120

Qy 121 MTENFETDGDQAQVYPNRFLPKWHSRINTFDEIEFIQLEKSTGKKGKTYPEIKA PW 180

Db 121 MTENFETDGDQAQVYPNRFLPKWHSRINTFDEIEFIQLEKSTGKKGKTYPEIKA PW 180

Qy 240 IAYTDWKETOE----KDPKGYVWNVNNDWMFKPG----AMAEVV-----KY-----AD 279

Db 202 MNPLWNSCSYEDSNQBEGPSTFDLISEFOALSRSRKAELVHFLLKTYAREPVTKAE 261

Qy 280 GWG-----GWWMLVNKEEKSXPDKNIVYPTPLKVELAQYNVNEVHPYPTVVKDALPEFTDVNQMY 336

Db 262 MUGSVVENVWQYFPVFSKASSSLQLVFGIELMEVDPIGHLYI-----PATCIGLSY 313

Qy 337 DALLNKSGATGVTFDPPTGVBL -KGISMDGGKAGVALVRSDYK-----TRDHKK 433

Db 314 DGLGDNQI-----MPKAGLILIVLAIARRGDCAPEKWIWBLSLVEPFEGREDISLG 367

Qy 391 STKLKLGFFHRARAMDGKAGVALVRSDYK-----TRDHKK 433

Db 368 DPKCLLJTOH-----FVQENYLERYQVPSDPACTEFLMGPRLVETSYVKV 413

Qy 434 GRHTARTSRSDYKFYD-----NRKIDSTSCHHHHHH 464

Db 414 LHHMVKITGGPHISYPPLHEWVIREGSETSGSHHHHHH 451

RESULT 10

ID AAY02348

XX AC AAY02348;

XX DT 17-OCT-2003 (revised)

DT 09-JUL-1999 (first entry)

DB Protein D, the fusion partner for HIV-1 Nef/rat fusion proteins.

XX XX HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;

XX vaccine; HIV infection; protein D.

XX OS WO940188-A2.

XX PN 12-AUG-1999.

XX PD 02-FEB-1999.

XX PF 99WO-EPO00660.

XX	98GB-00002543.	Bruck C,	Godart SAG,	Marchand M;
PR	98GB-00002543.	PI		
PR	98GB-00002543.	XX		
XX	98GB-00002543.	WPI; 1999-302282/25.		
XX		DR		
PA	(SMIK) SMITHKLINE BEECHAM BIOLOGICALS.	N-PSDB; AAX3588.		
XX		DR		
XX	Cabezon Silva T, Cohen J, Slaoui MM, Vinals Bassols C;	XX		
PI		PT		HIV Tat or Nef protein linked to a fusion partner.
XX		XX		
DR	WPI; 1999-494293/41.	PS Disclosure; Fig 2;	66pp;	English.
DR	N-PSDB; AAX8759J.	XX		
XX	New protein derivatives used in cancer vaccine therapy for treating a range of cancers including melanomas, carcinomas and cancers of breast.	The present sequence represents a fusion protein comprising LipD-HIV-1 NeF His. The protein is exemplifies the fusion proteins of the invention. The specification also describes fusion proteins comprising HIV-1 Tat protein. The fusion protein can be used in a vaccine to prevent HIV infection.		
PT		CC		
PT		CC		
XX		CC		
PS	Example 6; Page 67-68; 72pp; English.	CC		
XX	The present sequence represents a novel fusion protein composed of lipidated protein D (LipD) of Haemophilus influenzae B, the human MAGE-1 tumour-associated antigen and a hexahistidine tail. The invention relates to MAGE proteins fused to an immunological fusion partner such as LipD. The LipD moiety provides the fusion protein with additional exogenous T-cell epitopes and also increase expression levels in E. coli. The lipid cell tail ensures optimal presentation of the antigen to antigen-presenting cells. The affinity tag facilitates purification. The novel fusion proteins provide vaccines for immunotherapy of melanomas or other MAGE-associated tumours like breast, bladder, lung and non-small cell lung cancer, head and squamous cell carcinoma, colon carcinoma and oesophagus carcinoma.	Sequence 324 AA;		
CC		Query Match	24.5%;	Score 606; DB 2; Length 324;
CC		Best Local Similarity	37.3%;	Pred. No. 4.7e-44;
CC		Matches	168;	Mismatches 34; Indels 130;
CC		Gaps	15;	
Qy	19 CSSHSSNMANTQMKSDDKIIIAHGRSGYLPERTLLEQDLAMTKDGRL 78	Db	1 CSSHSSNMANTQMKSDDKIIIAHGRSGYLPERTLLEQDLAMTKDGRL 60.	
Db	79 VVTHDHFDLGLTDVAKKFPHRHKDGRYYVDFLTKEIOSLEMTEFNFTKDGKQAVQVYN 138	Qy	QV	
Db	61 VVTHHFHDGLTDVAKKFPHRHKDGRYYVDFLTKEIOSLEMTEFNFTKDGKQAVQVYN 138	Db	139 RFLPLKSHSPRIHTPDEIEFIGGLEKSTGKKVGTYPEIKAQWPHHIONQKDIAETLKVLK 198	
Qy	114 ---W-----SNSVYVG-WPTYRERMRRAPEAADDGVGAASRDLS 147	Qy	199 KYGTDRTDMVLYQTFDFNBLKR1KTELLPQMGMDLKLVQLIAVTWKEQTKDPKGYWV 258	
Db	Qy	148 KHGAIATSSNTAAATN-----AACAWLEAQEEBEVGFPV 179		
Db	259 NYNTDMFKPFGAMAEEVVKYADGVGFGWMLVNKEBKPDNIVTTPVKELAQXNVTEVHPY 318	Qy		
Qy	180 T-----PQVPLRPMVYKAADVDSHLFL---KEKGGLGEGLIHSQR-----214	Db	319 TVRKDALPEFTDYNQMYDALLNSGATGVFTDP---PDTGVBFELKGIKKSMDGCGKAGVA 374	
Db	Qy	215 -RQILDWVYHTO-----GYFPDQNYTGPGVRY----PLTRFCMCYKLV 255		
Db	Qy	375 LVRSYKLNKNSSNSTKLNLGHHRARAMDGKAGVALVRSDWKFYEDANGTRDHKKC 434		
Db	256 PVEPD-KVEEANKGENTSLLHPVSLH---GMDDPEREYLEWRFSRL-----AF 300	Qy		
Qy	435 RHTARTSRSDYKFYNDKR1DSTSGHzHHHHH 464	Db	301 HHVARBLHPBEY--FKN---CTSGHzHHHHH 324	
Db	RESULT 12			
AA	AY02352	ID	AAGG3234	RESULT 13
AC	AY02352 Standard; Protein: 324 AA.	ID	AAGG3234 standard; protein: 324 AA.	
XX		XX		
XX	DT 09-JUL-1999 (first entry)	XX		
DB	A representative LipD-Nef-His fusion protein.	XX		
XX	HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein; HIV vaccine; HIV infection; protein D.	XX		
XX	Synthetic.	XX		
OS	Human immunodeficiency virus 1.	XX		
XX	W0991684-A1.	XX		
XX	08-APR-1999.	XX		
XX	17-SEP-1998; 98WO-EP006040.	XX		
XX	26-SEP-1997; 97GB-00020585.	XX		
XX	(SMIK) SMITHKLINE BEECHAM BIOLOGICALS.	PA		
PA				

Db		301 HHVARELHPHY- -PKN-----CTSGHHHHHH 324
XX	PN	WO200154719-A2.
XX	PD	02-AUG-2001.
XX	PP	29-JAN-2001; 2001WO-EP000944.
XX	PR	31-JAN-2000; 2000GB-00002200.
PR	PR	14-APR-2000; 2000GB-00009336.
PR	PR	06-JUN-2000; 2000GB-000013806.
PR	PR	28-JUN-2000; 2000WO-EP005998.
XX	PA	(SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX	PI	Voss G;
XX	PS	WPI; 2001-476172/51.
DR	DR	N-PDBB; AAR42878.
XX	PT	New use of a human immunodeficiency virus (HIV) Tat, or Nef, or Nef protein linked to Tat (Nef-Tat) protein or polynucleotide and an HIV gp120 protein or polynucleotide for the manufacture of a vaccine.
XX	PT	Disclosure; Fig 1; 90pp; English.
XX	CC	The present sequence represents a His-tagged Nef protein of HIV, with a lipidation signal sequence (Lipod) which is removed after processing. The protein is expressed in Escherichia coli, and is used to produce the vaccine of the invention. The specification describes the use of HIV Tat, HIV Nef, or Nef-Tat, and HIV gp120 in the manufacture of a vaccine. The vaccine is used for the prophylactic or therapeutic immunization of humans against HIV. Tat, Nef or Nef-Tat act in synergy with gp120 in the treatment and prevention of HIV. The vaccine reduces the HIV viral load in HIV infected humans and results in a maintenance of CD4+ levels over those levels found in the absence of vaccination with HIV Tat, Nef or Nef -Tat and HIV gp120
XX	SQ	Sequence 324 AA;
Query Match	Best Local Similarity	24.5%; Score 606; DB 4; Length 324;
Matches 166; Conservative 34; Mismatches 118; Indels 130; Gaps 15;		
Qy	19	CSSHSSMANANTQMSDKIIIAHRGASGYLPENTLESKALAFAAQQADYLFDQLAMTKDGLR 78
Ddb	1	CSSHSSMANANTQMSDKIIIAHRGASGYLPENTLESKALAFAAQQADYLFDQLAMTKDGLR 60
Qy	79	VVIRDHFELDGLTDVAKFPHRHKGYYVDFLKEI0SLLEMENFETKDGRQAQVYN 138
Ddb	61	VVIRDHFELDGLTDVAKFPHRHKGYYVDFLKEI0SLLEMENFETMGKK----- 113
Qy	139	RFPWKSHPRHTPDEIEFIQGLELKSTGKVGTYPEKAPWFHQNGKDIAAETLYKVLR 198
Ddb	114	----W-----SKSSVVG-WPTYERMRDAEPAADGVGAASRDLE 147
Qy	199	KYGDKTDKDMVLQTDFNELKR1KTELLPQMGMIDLKLVQLIAYDWKETQDPKGYTV 258
Ddb	148	KHGALTSSNTAATN-----ACACWLEAQEEEBVGFPV 179
Qy	259	NYNTDWWMFKREGAMAEVVKYADGVGFYWLNKEESKPONIVTPVLAQTNVEVHPY 318
Ddb	180	T-----PQVPLRPMTYKAAVDLISHLFL--KEKGGLGLHSQR----- 214
Qy	319	TVRKDALPBFETTDYQNDMYDALLNKSGATGVFTDF---PDTGCVFGLKGTKSMDDGKAGVA 374
Ddb	215	--RQDIDWYHTQ-GYFDIQNYNTGPGVRY----PLTFRWCYLV 255
Qy	375	LVRSDYKLYNKNSNSNSTKLNGLBHRARAMDGKAGVALVRSDFKYEVDANGTRDHKGK 434
Ddb	256	PVERDVKVEANKGENTSLHHPVSLH--GMDDPERLEVLFWRPSL-----AF 300
Qy	435	RHTARTSRSDYKFDNKR1DSTSIGHHHHH 464
Db		RESULT 14 ADB67693 ID ADB67693 standard; protein; 324 AA. XX AC ADB67693; XX DT 04-DEC-2003 (first entry) XX DB HIV-1 Lipoprotein-His amino acid sequence SEQ ID NO:15. XX KW HIV; Tat; HIV-1; Nef; gp120; vaccine; immunisation; human; anti-HIV; XX KW fusion. XX OS Synthetic. OS Human immunodeficiency virus 1. XX PN WO2003011334-A1. XX PD 13-FEB-2003. XX PP 26-JUL-2002; 2002WO-EP008343. XX PR 27-JUL-2001; 2001GB-00018367. XX PA (GLAX) GLAXOSMITHKLINE BIOLOGICALS SA. PA (GLAX) GLAXO GROUP LTD. XX PI Ertel PF, Tite JP, Van Wely CA, Voss G; XX DR WPI; 2003-239474/23. DR N-PSDB; ADB67692. XX PT Use of an HIV Tat, Nef or gp120 protein or polynucleotide in the manufacture of a vaccine for a prime-boost delivery for the prophylactic or therapeutic immunization of humans against HIV. PT or therapeutic immunization of humans against HIV. XX Disclosure; Fig 1; 108pp; English. XX PS
Query Match	Best Local Similarity	24.5%; Score 606; DB 7; Length 324;
Matches 166; Conservative 34; Mismatches 118; Indels 130; Gaps 15;		
Qy	19	CSSHSSMANANTQMSDKIIIAHRGASGYLPENTLESKALAFAAQQADYLFDQLAMTKDGLR 78
Ddb	1	CSSHSSMANANTQMSDKIIIAHRGASGYLPENTLESKALAFAAQQADYLFDQLAMTKDGLR 60
Qy	79	VVIRDHFELDGLTDVAKFPHRHKGYYVDFLKEI0SLLEMENFETKDGRQAQVYN 138
Ddb	61	VVIRDHFELDGLTDVAKFPHRHKGYYVDFLKEI0SLLEMENFETMGKK----- 113
Qy	139	RFPWKSHPRHTPDEIEFIQGLELKSTGKVGTYPEKAPWFHQNGKDIAAETLYKVLR 198
Ddb	114	----W-----SKSSVVG-WPTYERMRDAEPAADGVGAASRDLE 147
Qy	199	KYGDKTDKDMVLQTDFNELKR1KTELLPQMGMIDLKLVQLIAYDWKETQDPKGYTV 258
Ddb	148	KHGALTSSNTAATN-----ACACWLEAQEEEBVGFPV 179
Qy	259	NYNTDWWMFKREGAMAEVVKYADGVGFYWLNKEESKPONIVTPVLAQTNVEVHPY 318
Ddb	180	T-----PQVPLRPMTYKAAVDLISHLFL--KEKGGLGLHSQR----- 214
Qy	319	TVRKDALPBFETTDYQNDMYDALLNKSGATGVFTDF---PDTGCVFGLKGTKSMDDGKAGVA 374
Ddb	215	--RQDIDWYHTQ-GYFDIQNYNTGPGVRY----PLTFRWCYLV 255
Qy	375	LVRSDYKLYNKNSNSNSTKLNGLBHRARAMDGKAGVALVRSDFKYEVDANGTRDHKGK 434
Ddb	256	PVERDVKVEANKGENTSLHHPVSLH--GMDDPERLEVLFWRPSL-----AF 300
Qy	435	RHTARTSRSDYKFDNKR1DSTSIGHHHHH 464
Db		Query Match
Best Local Similarity	37.3%; Pred. No. 4.7e-44;	
Matches 168; Conservative 34; Mismatches 118; Indels 130; Gaps 15;	Score 606; DB 7; Length 324;	
Qy	19	CSSHSSMANANTQMSDKIIIAHRGASGYLPENTLESKALAFAAQQADYLFDQLAMTKDGLR 78
Ddb	1	CSSHSSMANANTQMSDKIIIAHRGASGYLPENTLESKALAFAAQQADYLFDQLAMTKDGLR 60
Qy	79	VVIRDHFELDGLTDVAKFPHRHKGYYVDFLKEI0SLLEMENFETKDGRQAQVYN 138
Ddb	61	VVIRDHFELDGLTDVAKFPHRHKGYYVDFLKEI0SLLEMENFETMGKK----- 113
Qy	139	RFPWKSHPRHTPDEIEFIQGLELKSTGKVGTYPEKAPWFHQNGKDIAAETLYKVLR 198
Ddb	114	----W-----SKSSVVG-WPTYERMRDAEPAADGVGAASRDLE 147
Qy	199	KYGDKTDKDMVLQTDFNELKR1KTELLPQMGMIDLKLVQLIAYDWKETQDPKGYTV 258
Ddb	148	KHGALTSSNTAATN-----ACACWLEAQEEEBVGFPV 179
Qy	259	NYNTDWWMFKREGAMAEVVKYADGVGFYWLNKEESKPONIVTPVLAQTNVEVHPY 318
Ddb	180	T-----PQVPLRPMTYKAAVDLISHLFL--KEKGGLGLHSQR----- 214
Qy	319	TVRKDALPBFETTDYQNDMYDALLNKSGATGVFTDF---PDTGCVFGLKGTKSMDDGKAGVA 374
Ddb	215	--RQDIDWYHTQ-GYFDIQNYNTGPGVRY----PLTFRWCYLV 255
Qy	375	LVRSDYKLYNKNSNSNSTKLNGLBHRARAMDGKAGVALVRSDFKYEVDANGTRDHKGK 434
Ddb	256	PVERDVKVEANKGENTSLHHPVSLH--GMDDPERLEVLFWRPSL-----AF 300
Qy	435	RHTARTSRSDYKFDNKR1DSTSIGHHHHH 464

Qy 79 VVIIHDHFLDGLTDVAKKPCHRHKCGRYVVDPTLKBHQISLEMNTNPFETKDGKQAOQVYPN 138
 Db 61 VVIIHDHFLDGLTDVAKKPCHRHKCGRYVVDPTLKBHQISLEMNTNPFETKDGK----- 113
 Qy 139 RFPFLWKSFRHTFEDIEFQGLEKSTGKGVYYPEIRKAQFWFHONGKDIAAETLKVKR 198
 Db 114 ----W----- SKSSYVG-NPTVREMRRAEPADVGVAASRDLE 147
 Qy 199 KYGYDKKTMVYLQTEDEFNELKRIKTELLPQMGMILKLVQIAYTDWKETQEKDPKGMYV 258
 Db 148 KHGATTSNTAATN-----AACAWLEAQEEBEGVGPV 179
 Qy 259 NYNDMMEPKGAMAEVVKYADGVGPWMVLVNEKESSPKDNIVYTPLVKELAQYNVEVHPY 318
 Db 180 T-----POVPLRPMTYKAADVLSHEL-----KEKGGLGEIHSOR----- 214
 Qy 319 TVRKDALPEFTIDVNOMYDALLNKSGATGVFTDF----PDIGVEFLKGTKSMDDGCRAGVA 374
 Db 215 -RQDILDWLWHTQ-----GYFPDWQNYTPGPGRVY----PLTFGMCYKLV 255
 Qy 375 LVRSDYKLYNKNSSNSTKLNLGEHRARAMGGKAGVALVRSDFKEYEDANGTORDHKG 434
 Db 256 PVEPD-KVEEANKGENTSLHPVSLH---GMDDPEREVLWRFDTRL-----AP 300
 Qy 435 RHTARTSRSDYKFDYDKRIBDTSGHHHHH 464
 Db 301 HHVARELHPEY-FKN---CTSGHHHHH 324
 RESULT 15
 ID. AAY02354 standard; protein; 326 AA.
 XX .
 DT 09-JUL-1999 (first entry)
 DE A representative LipoD-Nef fusion protein.
 XX HIV nef gene; fusion protein; HIV nef gene; Nef protein; Rat protein;
 XX vaccine; HIV infection; protein D.
 OS Synthetic.
 XX Human immunodeficiency virus 1.
 PN WO9916884.A1.
 XX 08-APR-1999.
 XX 17-SEP-1999; 98WO-BP006040.
 XX PR 26-SEP-1997; 97GB-00020585.
 XX PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX PI Bruck C, Godart SAG, Marchand M;
 XX DR WPI; 1999-302282/25.
 XX DR N-PSDB; AAX35690.
 PT HIV Tat or Nef protein linked to a fusion partner.
 PS Disclosure: Fig 2: 66pp; English.

XX The present sequence represents a fusion protein comprising LipoD-HIV-1
 CC Nef. The protein is exemplifies the fusion proteins of the invention. The
 CC specification also describes fusion proteins comprising HIV-1 Tat
 CC protein. The fusion protein can be used in a vaccine to prevent HIV
 CC infection.
 XX Sequence 326 AA;
 PS Query Match 24.1%; Score 597; DB 2; Length 326;

Best Local Similarity 37.2%; Pred. No. 2.9e-43;
 Matches 167; Conservative 34; Mismatches 118;
 Indels 130; Gaps 15;

Qy 20 SSHSSMMANTQMSDKITIAHRGSGYLPDETHLESKALAFAQOQDYLEQDLAMTKDGRLV 79
 Db 4 SSHSSMMANTQMSDKITIAHRGSGYLPDETHLESKALAFAQOQDYLEQDLAMTKDGRLV 63
 Qy 80 VHDFIFLDGTTDVAKKFPHRHDKDRGYYVDFTURERIQLSELEMTEFETKDGKQAOQVYPNR 139
 Db 64 VHDFIFLDGTTDVAKKFPHRHDKDRGYYVDFTURERIQLSELEMTEFETKDGKQAOQVYPNR 115
 Qy 140 FPLWKSFRHTFEDIEFQGLEKSTGKGVYTYPEIKAPWFHONGKDIAAETLKVKR 199
 Db 116 ---W-----SKSSYVG-WPTVRRMRAAPAAADGVGAASRDLEX 150
 Qy 200 YGYDKKTDWVYLQTPDFENELKRIKTELLQMGMDLKLVOLIAVTDWKEQDPKGYWNT 259
 Db 151 HGAITSNTAATN-----AACAWLEAQEEEEVGFPTV 182
 Qy 260 YNYDMFMFKGAMAEVVKYADGVGPWYMLVNKEESSPKDNIVYTPLVKELAQYNVEVHPY 319
 Db 183 -----POVPLRPMTYKAADVLSHEL-----KEKGGLGEIHSOR----- 216
 Qy 320 VRKDALPEFTIDVNQMYDALLNKSGATGVFTDF----PDIGVEFLKGTKSMDDGGRAGVAL 375
 Db 217 -RQDILDWLWHTQ-----GYFPDWQNYTPGPGRVY----PLTFGMCYKLV 258
 Qy 376 VRSDYKLYNKNSSNSTKLNLGEHRARAMGGKAGVALVRSDFKEYEDANGTORDHKG 435
 Db 259 VEPD-KVEEANKGENTSLHPVSLH---GMDDPEREVLWRFDTRL-----APH 303
 Qy 436 HTARTSRSDYKFDYDKRIBDTSGHHHHH 464
 Db 304 HVARELHPEY-FKN---CTSGHHHHH 326

Search completed: June 22, 2005, 11:37:45
 Job time : 166 secs

Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 22, 2005, 11:32:20 ; Search time 43 Seconds (without alignments)

805.515 Million cell updates/sec

Title: US-09-719-379A-81

Perfect score: 2477

Sequence: 1 MDPKTLAISLLAAAGVLAGCS.....YKFYDNKRIDSSTGHHHHH 464

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database : Issued Patents AA:
 1: /cgn2_6/picodata/1/iaa/5A_COMB.pep:
 2: /cgn2_6/picodata/1/iaa/5B_COMB.pep:
 3: /cgn2_6/picodata/1/iaa/6A_COMB.pep:
 4: /cgn2_6/picodata/1/iaa/6B_COMB.pep:
 5: /cgn2_6/picodata/1/iaa/8C_PCTUS_COMB.pop:
 6: /cgn2_6/picodata/1/iaa/backfiles.pop:
 * Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	1251.5	50.5	359	4	US-09-489-039A-8365	Sequence 8365, AP
2	1219.5	49.2	364	4	US-09-543-68IA-6003	Sequence 6003, AP
3	565	22.8	220	3	US-09-485-885-1	Sequence 1, Appli
4	565	22.8	220	3	US-09-485-885-8	Sequence 8, Appli
5	563	22.7	278	3	US-09-485-885-21	Sequence 21, Appli
6	561.5	22.7	227	3	US-09-485-885-19	Sequence 19, Appli
7	559.5	22.6	273	3	US-09-485-885-4	Sequence 4, Appli
8	559.5	22.6	371	3	US-09-485-885-6	Sequence 6, Appli
9	558	22.5	227	3	US-09-485-885-23	Sequence 16, Appli
10	558	22.5	383	4	US-09-485-885-23	Sequence 23, Appli
11	381.5	15.4	386	4	US-09-328-352-5514	Sequence 5514, AP
12	381	15.4	503	4	US-09-552-991A-27882	Sequence 27882, A
13	269.5	10.9	316	4	US-08-956-171E-5331	Sequence 5231, AP
14	269.5	10.9	316	4	US-08-951-184-2	Sequence 5231, AP
15	265	10.7	353	3	US-09-134-001C-3115	Sequence 3115, AP
16	245.5	9.9	383	4	US-09-252-991A-25535	Sequence 25535, AP
17	201	8.1	260	3	US-09-134-001C-5097	Sequence 5097, AP
18	182.5	7.4	359	1	US-08-457-991B-2	Sequence 2, Appli
19	182.5	7.4	359	3	US-08-467-722A-2	Sequence 2, Appli
20	182.5	7.4	359	4	US-09-451-184-2	Sequence 2, Appli
21	173	7.0	247	4	US-09-107-532A-3856	Sequence 3856, AP
22	169.5	6.8	203	4	US-09-210-1794	Sequence 1794, AP
23	149.5	6.0	338	1	US-08-210-394-1	Sequence 1, Appli
24	148.5	6.0	256	3	US-09-134-001C-4444	Sequence 4544, AP
25	143.5	5.8	318	4	US-09-107-532A-6446	Sequence 6446, AP
26	133	5.4	382	4	US-09-134-000C-3229	Sequence 3829, AP
27	130.5	5.3	256	4	US-09-583-110-5215	Sequence 5215, AP

ALIGNMENTS

RESULT 1
 US-09-489-039A-8365
 ; Sequence 8365, Application US/094890909A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; FILE OF INVENTION: PNEUMONIA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709 2004001
 ; CURRENT APPLICATION NUMBER: US/09/449,039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO: 8365
 ; LENGTH: 359
 ; TYPE: PRT
 ; ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-8365

Query Match 50.5%; Score 1251.5; DB 4; Length 359;
 Best Local Similarity 65.3%; Pred. No. 4.2e-108; M1 matches 48; Indels 1; Gaps 1;
 Matches 231; Conservative 7 ALSLAAAGVLAGCSSHSSMNANTQNSDKTIIAHGASGYLPENTLESKALAPAQADYL 66
 Qy 3 AMKKMLTALMGMLSSALCFSATADKVNIAHGA SGVLPENTLPAKAMAYAQCADYL 62
 Db 67 EQDLAMTKDGRLLVWHDHFGLGLTVAKKEPFRHRKDGRYVYDITFLKEQSLENTENF 126
 Qy 63 EQDLAMTKDGRLLVWHDHFGLGLTVAKKEPFRHRKDGRYVYDITFLDEIKSLKFTEGFE 122
 Db 127 TKDGKOAQVYYPNRPLWKSFRHTFEDSLEBFIQGLEKSTGKKGVGYPEBKAPWPHQNG 186
 Qy 123 PKNGQVQTYPGRFMGKSDFRIMTFEEETBFVQGHNSTGKNIQYBPKIAPWPHQEG 182
 Db 187 KDIATETLKVLYKKGYDKKTDMVYLTQPDNLXRIKTELLPQMGMDLKLYQLIAYTDWK 246
 Qy 183 KDIANSTLKVLYKEYGTSKQDKVYVQCFDANEKLRIKINLEPKGMNDLNVLQLIAYTDWN 242
 Db 247 ETQBDPKGTVWNTYDWMFKPGANAEVVYADGVPGPQVFTDPDNIVTPPLVK 306
 Qy 243 ETQQROADGKVNNTSYDWMFKPGANAQIAQYADGVPDTMLV-LEGSKGCAVKLTAMVK 301
 Db 307 ELAQVNVEHPYTVRDALPEFFTVYDQKGVFTFPDTGVEFL 360
 Qy 302 EAHASHLQLVHPYTVRADOLPEYATVNVQLYDGLFTDFDKAVQFL 355

US-09-543-681A-6003 ; Sequence 6003; Application US/09543681A
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709-1002-001
; CURRENT APPLICATION NUMBER: US/09/543, 681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128, 706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO: 6003
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Proteus mirabilis
; US-09-543-681A-6003

Query Match 49.2%; Score 1219.5; DB 4; Length 364;
Best Local Similarity 64.7%; Pred. No. 4.1e-105; Indels 5; Gaps 3;
Matches 229; Conservative 50; Mismatches 70; Indels 5; Gaps 3;

RESULT 4
US-09-485-885-8
; Sequence 8; Application US/09485885
; Patent No. 642224
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; Cabezon, Silvia, Teresa
; Delisse, Anne-Marie Eva Fernande
; Gerard, Catherine Marie Ghislaine
; Ghislaine
; APPARTANT: Lombardo-Bencheikh, Angela
; TITLE OF INVENTION: Vaccine
; FILE REFERENCE: B45107
; CURRENT APPLICATION NUMBER: US/09/485, 885
; CURRENT FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: PCT/EP98/05285
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: GB 9117953.5
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ For Windows Version 3.0
; SEQ ID NO: 8
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Homo sapien
; US-09-485-885-8

Query Match 22.8%; Score 565; DB 3; Length 220;
Best Local Similarity 68.0%; Pred. No. 1.7e-44; Indels 8; Gaps 2;
Matches 117; Conservative 3; Mismatches 3; Indels 4; Gaps 2;

Query 20 SSHSSNNANTQMSDKITIAHARGASGYLPETHTLESKALAFAAQQADYLEDQLAMTKDGRLV 79
Db 4 SSHSSNNANTQMSDKITIAHARGASGYLPETHTLESKALAFAAQQADYLEDQLAMTKDGRLV 63

Qy 80 VHDFDLDGITYDVAKKFPHRHKDRGRRYVDFTLIKEIQLSLEMENFETK----- 128
Db 64 VHDFDLDGITYDVAKKFPHRHKDRGRRYVDFTLIKEIQLSLEMENFETMAMHGDTPTLHE 123

Qy 129 -----DGKQAOQVYPNRFPLWKSFRHTF 152
Db 124 YMIDLQPETTDLYCQEQLNDSSEREDEIDPGAQAEPRD----AHYNIVTP 170

RESULT 5
US-09-485-885-21
; Sequence 21; Application US/09485885
; Patent No. 642224
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; Cabezon, Silvia, Teresa
; Delisse, Anne-Marie Eva Fernande
; Gerard, Catherine Marie Ghislaine
; Ghislaine
; APPARTANT: Lombardo-Bencheikh, Angela
; TITLE OF INVENTION: Vaccine
; FILE REFERENCE: B45107

Query Match 22.8%; Score 565; DB 3; Length 220;
Best Local Similarity 68.0%; Pred. No. 1.7e-44; Indels 8; Gaps 2;
Matches 117; Conservative 3; Mismatches 3; Indels 4; Gaps 2;

Query 20 SSHSSNNANTQMSDKITIAHARGASGYLPETHTLESKALAFAAQQADYLEDQLAMTKDGRLV 79
Db 4 SSHSSNNANTQMSDKITIAHARGASGYLPETHTLESKALAFAAQQADYLEDQLAMTKDGRLV 63

Qy 80 VHDFDLDGITYDVAKKFPHRHKDRGRRYVDFTLIKEIQLSLEMENFETK----- 128
Db 64 VHDFDLDGITYDVAKKFPHRHKDRGRRYVDFTLIKEIQLSLEMENFETMAMHGDTPTLHE 123

Qy 129 -----DGKQAOQVYPNRFPLWKSFRHTF 152
Db 124 YMIDLQPETTDLYCQEQLNDSSEREDEIDPGAQAEPRD----AHYNIVTP 170

RESULT 3
US-09-485-885-1
; Sequence 1; Application US/09485885
; Patent No. 634224
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; Cabezon, Silvia, Teresa
; Delisse, Anne-Marie Eva Fernande
; Gerard, Catherine Marie Ghislaine
; Ghislaine
; APPARTANT: Lombardo-Bencheikh, Angela
; TITLE OF INVENTION: Vaccine
; FILE REFERENCE: B45107

CURRENT APPLICATION NUMBER: US/09/485, 885
CURRENT FILING DATE: 2000-02-18
PRIOR APPLICATION NUMBER: PCT/EP98/05285
PRIOR FILING DATE: 1998-08-17
PRIOR APPLICATION NUMBER: GB 9117953.5
PRIOR FILING DATE: 1997-08-22
NUMBER OF SEQ ID NOS: 23
SOFTWARE: FastSEQ For Windows Version 3.0
SEQ ID NO: 1
LENGTH: 220
TYPE: PRT
ORGANISM: Homo sapien
US-09-485-885-1

CURRENT APPLICATION NUMBER: US/09/485,885
 CURRENT FILING DATE: 2000-02-18
 PRIOR APPLICATION NUMBER: PCT/EP98/05285
 PRIOR FILING DATE: 1998-08-17
 PRIORITY NUMBER: GB 9717953.5
 NUMBER OF SEQ ID NOS: 23
 SOFTWARE: FastSEQ for Windows Version 3.0
 SEQ ID NO: 21
 TYPE: PRT
 ORGANISM: Homo sapien

Query Match Score 22.7%; DB 3; Length 27;
 Best Local Similarity 34.5%; Pred. No. 3.8e-44;
 Matches 157; Conservative 29; Mismatches 79; Indels 190; Gaps 13;

Qy 20 SSSHSSNANTQMSKDKIIIAHRCAGYLPEHTLESKALAPQQADYLEQDLMTKDGRLL 79
 Db 4 SSSHSSNANTQMSKDKIIIAHRCAGYLPEHTLESKALAPQQADYLEQDLMTKDGRLL 63

Query Match Score 563; DB 3; Length 27;
 Best Local Similarity 34.5%; Pred. No. 3.8e-44;
 Matches 157; Conservative 29; Mismatches 79; Indels 190; Gaps 13;

Qy 20 SSSHSSNANTQMSKDKIIIAHRCAGYLPEHTLESKALAPQQADYLEQDLMTKDGRLL 79
 Db 4 SSSHSSNANTQMSKDKIIIAHRCAGYLPEHTLESKALAPQQADYLEQDLMTKDGRLL 63

RESULT 7
 US-09-485-885-4
 Sequence 4, Application US/09485885
 ; Patent No. 6342224
 ; GENERAL INFORMATION:
 ; APPLICANT: Bruck, Claudine
 ; APPLICANT: Cabeson, Silvia, Teresa
 ; APPLICANT: Delisse, Anne-Marie Eva Fernande
 ; APPLICANT: Gerard, Catherine Marie Ghislaine
 ; APPLICANT: Lombardo-Bancheikh, Angelia
 ; TITLE OF INVENTION: Vaccine
 ; FILE REFERENCE: B451.07
 ; CURRENT APPLICATION NUMBER: US/09/485,885
 ; CURRENT FILING DATE: 2000-02-18
 ; PRIOR APPLICATION NUMBER: PCT/EP98/05285
 ; PRIOR FILING DATE: 1998-08-17
 ; PRIOR APPLICATION NUMBER: GB 9717953.5
 ; PRIOR FILING DATE: 1997-08-22
 ; NUMBER OF SEQ ID NOS: 23
 ; SEQ ID NO: 4
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; LENGTH: 273
 ; TYPE: PRT
 ; ORGANISM: Homo sapien

US-09-485-885-4

Query Match Score 559.5; DB 3; Length 273;
 Best Local Similarity 91.1%; Pred. No. 7.8e-44;
 Matches 112; Conservative 1; Mismatches 9; Indels 1; Gaps 1;

Qy 20 SSSHSSNANTQMSKDKIIIAHRCAGYLPEHTLESKALAPQQADYLEQDLMTKDGRLL 79
 Db 4 SSSHSSNANTQMSKDKIIIAHRCAGYLPEHTLESKALAPQQADYLEQDLMTKDGRLL 63

Query Match Score 563; DB 3; Length 273;
 Best Local Similarity 91.1%; Pred. No. 7.8e-44;
 Matches 112; Conservative 1; Mismatches 9; Indels 1; Gaps 1;

Qy 20 SSSHSSNANTQMSKDKIIIAHRCAGYLPEHTLESKALAPQQADYLEQDLMTKDGRLL 79
 Db 4 SSSHSSNANTQMSKDKIIIAHRCAGYLPEHTLESKALAPQQADYLEQDLMTKDGRLL 63

RESULT 8
 US-09-485-885-6
 Sequence 6, Application US/09485885
 ; Patent No. 6342224
 ; GENERAL INFORMATION:
 ; APPLICANT: Bruck, Claudine
 ; APPLICANT: Cabeson, Silvia, Teresa
 ; APPLICANT: Delisse, Anne-Marie Eva Fernande
 ; APPLICANT: Gerard, Catherine Marie Ghislaine
 ; APPLICANT: Lombardo-Bancheikh, Angelia
 ; TITLE OF INVENTION: Vaccine
 ; FILE REFERENCE: B451.07
 ; CURRENT APPLICATION NUMBER: US/09/485,885
 ; CURRENT FILING DATE: 2000-02-18
 ; PRIOR APPLICATION NUMBER: PCT/EP98/05285
 ; PRIOR FILING DATE: 1998-08-17
 ; PRIOR APPLICATION NUMBER: GB 9717953.5
 ; PRIOR FILING DATE: 1997-08-22
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO: 19
 ; LENGTH: 227

RESULT 10
 US-09-485-885-23
 ; Sequence 23, Application US/09485885
 ; Patent No. 634224
 ; GENERAL INFORMATION:
 ; APPLICANT: Bruck, Claudine
 ; APPLICANT: Cabezon Silva, Teresa
 ; APPLICANT: Delisse, Anne-Marie Eva Fernande
 ; APPLICANT: Gerard, Catherine Marie Ghislaine
 ; APPLICANT: Lombardo-Bencheikh, Angela
 ; TITLE OF INVENTION: Vaccine
 ; FILE REFERENCE: B45107
 ; CURRENT APPLICATION NUMBER: PCT/EP98/05285
 ; CURRENT FILING DATE: 2000-02-18
 ; PRIOR APPLICATION NUMBER: PCT/EP98/05285
 ; PRIOR FILING DATE: 1998-08-17
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO: 6
 ; LENGTH: 371
 ; TYPE: PRT
 ; ORGANISM: Homo sapien
 ; US-09-485-885-6

Query Match 22.6%; Score 559.5; DB 3; Length 371;
 Best Local Similarity 91.1%; Pred. No. 1.2e-43;
 Matches 112; Conservative 1; Gaps 1;
 Indels 1; Mismatches 9;

Qy 20 SSSSSNNTQNSKDKIIIAHRGASGYLPENTLESKALAFACQADYLEQDLAMTKDGRLV 79
 Db 4 SSSSSNNTQNSKDKIIIAHRGASGYLPENTLESKALAFACQADYLEQDLAMTKDGRLV 63

Qy 80 VHDHFGLDGLTDVAKKFPHRKDGYYVIDFTLKEIQSLENTEFNPKDGKQ-AQYVNP 138
 Db 64 VHDHFGLDGLTDVAKKFPHRKDGYYVIDFTLKEIQSLENTEFNPKDQPERPR 123

Qy 139 RFP 141
 Db * 124 XLP 126

RESULT 9
 ; Sequence 16, Application US/09485885
 ; Patent No. 634224
 ; GENERAL INFORMATION:
 ; APPLICANT: Bruck, Claudine
 ; APPLICANT: Cabezon Silva, Teresa
 ; APPLICANT: Delisse, Anne-Marie Eva Fernande
 ; APPLICANT: Gerard, Catherine Marie Ghislaine
 ; APPLICANT: Lombardo-Bencheikh, Angela
 ; TITLE OF INVENTION: Vaccine
 ; FILE REFERENCE: B45107
 ; CURRENT APPLICATION NUMBER: US/09/485-885
 ; CURRENT FILING DATE: 2000-02-18
 ; PRIOR APPLICATION NUMBER: PCT/EP98/05285
 ; PRIOR FILING DATE: 1998-08-17
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO: 16
 ; LENGTH: 227
 ; TYPE: PRT
 ; ORGANISM: Homo sapien
 ; US-09-485-885-16

Query Match 22.5%; Score 558; DB 3; Length 227;
 Best Local Similarity 100.0%; Pred. No. 8.1e-44;
 Matches 108; Conservative 0; Gaps 0;
 Indels 0; Mismatches 0;

Qy 20 SSSSSNNTQNSKDKIIIAHRGASGYLPENTLESKALAFACQADYLEQDLAMTKDGRLV 79
 Db 4 SSSSSNNTQNSKDKIIIAHRGASGYLPENTLESKALAFACQADYLEQDLAMTKDGRLV 63

Qy 80 VHDHFGLDGLTDVAKKFPHRKDGYYVIDFTLKEIQSLENTEFNPKDQPERPR 127
 Db 64 VHDHFGLDGLTDVAKKFPHRKDGYYVIDFTLKEIQSLENTEFNPKDQPERPR 111

RESULT 11
 ; Sequence 11, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETE^N NI FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC99-03PA
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; CURRENT FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO: 5514
 ; LENGTH: 386
 ; TYPE: PRT
 ; ORGANISM: Acinetobacter baumannii
 ; US-09-328-352-5514

Query Match 15.4%; Score 381.5; DB 4; Length 386;
 Best Local Similarity 31.6%; Pred. No. 5e-27;
 Matches 125; Conservative 51; Gaps 75;
 Indels 145; Mismatches 145;

Qy 6 LALSLIAAGVLAGCSSH---SNNNTQNSKDKIIIAHRGASGYLPENTLESKALAFACQADYLEQDLAMTKDGRLV 61
 Db 13 ICLSLIS---LVGNDDDKDTETTPPEYQLPKLIVYGRGASALRPNTLASYQKAIDD 69

Qy 62 QADYLEQDLAMTKDGRLVIDFTLKEIQSLENTEFNPKDQPERPR 127
 Db 70 GADFLEPDLYSTKDGVLVARNENE GGTTNTVSTLSQFADRKTKTNDGVDTGWWTBDFT 129

Query Match 22.5%; Score 558; DB 3; Length 227;
 Best Local Similarity 100.0%; Pred. No. 8.1e-44;
 Matches 108; Conservative 0; Gaps 0;
 Indels 0; Mismatches 0;

Qy 20 SSSSSNNTQNSKDKIIIAHRGASGYLPENTLESKALAFACQADYLEQDLAMTKDGRLV 79
 Db 4 SSSSSNNTQNSKDKIIIAHRGASGYLPENTLESKALAFACQADYLEQDLAMTKDGRLV 63

Qy 80 VHDHFGLDGLTDVAKKFPHRKDGYYVIDFTLKEIQSLENTEFNPKDQPERPR 127
 Db 64 VHDHFGLDGLTDVAKKFPHRKDGYYVIDFTLKEIQSLENTEFNPKDQPERPR 111

Query 170 VGIYPEIKAP-WFHONGKDIAB--TLKVLLKKGYDRTDMVYLOTFDFNLKRKTEL 226
 Database 179 IGHYIETKHPYFKQNQ--LAMEDTLIKTLAKYCYTRDIAPVYLQSPEVNLKDILKREL 235

Query 227 LPQMGMDLKLKVOLIAYTWDKETQEKDPKGWYNNYNDWMFKEGAMAEVVKYADGVGP-G 284
 Database 236 -DLHKTKLKHQIQLYDVKTSRPADEVESGDTKT'ADLATAQGLKDVKYANGVGPSPKG 293

Query 285 WYMLVNKEBESKDNNIVYTPLVKELAQYNTVEHPTVYRKD----- 323
 Database 294 YLTENNDGSYKTC----STFSDAHTAGLKTHPYTPEPNFLPAPLKCSPDKPAERCPT 349

Query 324 -ALPEFTDVNOMYDALLINKSAGTAGVTFDPPTGVE 358
 Database 350 GALKEF----EAY---FKAGVDGVTFDDPALGRE 376

RESULT 12
 Sequence 27882, Application US/09252991A.
 GENERAL INFORMATION:
 APPLICANT: Marc J. Rubenfield et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 FILE REFERENCE: 107196.136
 CURRENT APPLICATION NUMBER: US/09/252,911A
 CURRENT FILING DATE: 1999-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO: 27882
 LENGTH: 503
 TYPE: PRT
 ORGANISM: Pseudomonas aeruginosa
 US#09-252-991A-27882

Query Match 15.4%; Score 381; DB 4; Length 503;
 Best Local Similarity 30.5%; Pred. No. 8.4e-7; Matches 119; Conservative 64; Gaps 13;
 Matches 119; Conservatve 48; Mismatches 116; Indels 107; Gaps 16;

Query 7 ASSLARGVLAGCSSHSSNM-----ANTOM-----KSDKIIIAHGRGASGYLPHEHTL 52
 Database 133 ACMLMAGLASAHADTGNCVRAIDNAQGDSVRHPAGHERSPLVTAHGRGASGYPHEHTL 192
 Query 53 ESKALAPAQQAQYALEQDLMITKDGRLVVTHDHLGDTDVAK--KPHRARK--DG-- 104
 Database 193 GAYALAVMMGADYEPDVLMTRDGKLYVARHNBELGLTVAQHPEFADRKTKQVGDVEL 252

Query 105 -RYVYDFTLKEIISLQEMENTFBTKDQKAQVYPNRFPLWKSHFRHTFEDIEFQGLE 163
 Database 253 TGWSFEDTLEBLKTRAIERPT-----IRPGNARL-DGTFFPILTQETIDLVKSLO 304

Query 164 KSTGKKVGKYPEKAPKPHONGKDIAAETLKVLLKGYDKTKDMVYLOTFDFNLKRK 223
 Database 305 ISQORTIGLYPKIGHFTHFORGLAMERPLVTLHNGYLGPRAFYFIQSFPEVNILK 364

Query 224 TELLPQMDLKLVQLIATWDKETQEKDPKGWYNNYNDWMFKEGAMAEVVKYADGVGP 283
 Database 365 RL-----TGIRLVLQ-----YGSQPYDQQAAGGSUTYAEMATAKLRQVARYAYGVGP 413

Query 284 GYMLVNKEBESKDNNIVYTPLVKELAQYNTVEHPTVYRKD----- 327
 Database 414 DKSYVIPRDANGNLGP-----TRFVDAHAAGLKVHPTFRAENSFLPAERSADGNPQ 468

Query 328 FFTDVNOMYDALLINKSAGTAGVTFDPPTGVE 357
 Database 469 SRGDLACEIRAYLD-ACTDGL-FSDQPVAV 497

US#09-956-171E-5231
 Sequence 5231, Application US/0956171E
 Patent No. 6593114
 GENERAL INFORMATION:
 APPLICANT: Charles Runsch
 Gil H. Choi
 Patrick S. Dillon
 Craig A. Rosen
 Steven C. Barash
 Michael R. Farnon
 ADDRESS: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: Maryland
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MSDOS version 6.2
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/956,171E
 FILING DATE: 20-Oct-1997
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/009,861
 FILING DATE: January 5, 1996
 APPLICATION NUMBER: 08/781,986
 FILING DATE: January 3, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Mark J. Hyman
 REGISTRATION NUMBER: 46,789
 REFERENCE/DOCKET NUMBER: PB248P1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (240) 314-1224
 TELEFAX: (301) 309-8439
 INFORMATION FOR SEQ ID NO: 5231:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 316 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 5231:
 US#09-956-171E-5231

Query Match 10.9%; Score 269.5; DB 4; Length 316;
 Best Local Similarity 27.3%; Pred. No. 1e-16; Mismatches 116; Indels 107; Gaps 16;
 Matches 103; Conservative 48; Mismatches 116; Indels 107; Gaps 16;

Query 8 LSLLAAGVLAGCSSHSSNM-----TOMKSKI-IIAHGRGASGYLPHEHTLBSKALA 58
 Database 30 LSVPATG-----AEQTNQIANKPQIAQWHTNLNTERPTTIAHGRGASGYAEPHTFQYDKS 84

Query 59 FAQ-QADYLEQDLMATKDGRLUVVTHDHFQDGLTDVAKFPHRHKDGRGYVVDFTLKEIQ 117
 Database 85 HNELKASYIETDQTKDGHLVAMDETVVRT-----NGRKVEDYTDDELK 132

Query 118 SLEMTEFNFTDKGKQAVQ-YPNRFPLWKSHFRHTFEDIEFQGLEKSKKKVGYPEI 176
 Database 133 QLDAGSSWENKCKYKVARASTYNA-----KVPTLDELE-----RGPNANYIET 177

Query 177 KAPWWHHQNGKDIAAETLKVLY-----GVDKKTQVYLQFDFBLKRKTFTELLPQGMG 232
 Database 178 KSPDVY----PGMEBOLLASLKKHILLNNNKKLKHMIQSFSDELSLKKHRQ----- 226

Query 233 DLKLYOLIAYTDWKETQEKDPKGWYNNYNDWMFKEGAMAEVVKYADGVGPWMLVNKE 292
 Database 227 -NKYJBLVLUVDKGELQQ-----ENDQRLKEIRSTAIGLGD----- 262

RESULT 14
 US-08-781-986A-5231
 ; Sequence 5231, Application US/08781986A
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles Kunsch
 ; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
 ; NUMBER OF SEQUENCES: 5255
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEES: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 ; COMPUTER: HP Vectra 486/33
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; CURRENT APPLICATION NUMBER: US/08/781,986A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Benson, Bob
 ; REGISTRATION NUMBER: 30,446
 ; TELECOMMUNICATION INFORMATION: PR248PP
 ; TELEPHONE: (301) 309-8504
 ; TELEFAX: (301) 309-8512
 ; INFORMATION FOR SEQ ID NO: 5231:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 316 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-781-986A-5231

Query 293 ESKPDNNIVTPLVKE---LAQVNVEVHPYTVRKDALPEFTDYNQMDAL-LNKGATG 347
 Db 263 -----YDLTEQNTTHILKDLGFIYHPT-----WNEADMRLRINKYGVDG 302
 Qy 348 VFTDFPDTGVEFLK 361
 Db 303 VFTNFADKYKEVIK 316

Query 233 DLKLVQLIAYTDWKETOEKDPKGYWWVNTNYDWMFKPGAMAEVVKYADGVGPQNMVLVNKE 292
 Db 227 -NKGTPVPLVLDKCBLOO-----FNDRQKEITSYAIGPDI----- 262
 Qy 293 ESKPDNNIVTPLVKE---LAQVNVEVHPYTVRKDALPEFTDYNQMDAL-LNKGATG 347
 Db 263 -----YDLTEQNTTHILKDLGFIYHPT-----WNEADMRLRINKYGVDG 302
 Qy 348 VFTDFPDTGVEFLK 361
 Db 303 VFTNFADKYKEVIK 316

RESULT 15
 US-09-134-001C-3115
 ; Sequence 3115, Application US/09134001C
 ; PATENT NO. 6380310
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 ; TITLE OF INVENTION: EPIDERMIDS FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC-007
 ; CURRENT APPLICATION NUMBER: US/09/134,001C
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/064,964
 ; PRIOR FILING DATE: 1997-11-08
 ; PRIOR APPLICATION NUMBER: US 60/055,779
 ; PRIOR FILING DATE: 1997-08-14
 ; NUMBER OF SEQ ID NOS: 5674
 ; SEQ ID NO: 3115
 ; LENGTH: 353
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus epidermidis
 ; US-09-134-001C-3115

Query Match 10.9%; Score 269.5; DB 4; Length 316;
 Best Local Similarity 27.5%; Pred No. 1e-16; Mismatches 116; Indels 107; Gaps 16;

Query 5 TLAISLAAVGVLAGCGSSHS-----SNN----ANT---OMKSPIKI--- 36
 Db 20 TGLASLSSPSIAETSNHSPAPNSSEKSQQQTINIPSNMNOSSAKTFKQQAESDSVSQGT 79
 Qy 37 -----1IAHRGASGYLPEHTLESKALFAQ-QADYLEBODLAMTKDGRLL 78
 Db 80 NDTHNHWTKNLTGERFTTAHRGASGYAPEHTFAYDKSHNELGASYIEIDLQRTDGH 139
 Qy 79 VVIHPHFQDGLTDVAKKEPHRHKDGRYYVIDFTKETOSLEMENFTKGKQAOVYPN 138
 Db 140 VAMHDEKVNRTT-----NGHGRVDQLTKEKLQLDGWSWNRKHPEAK--N 184
 Qy 139 RFPWKSHPPIIHTPDELETFIQGIKSTKVGTYPEIKAPWFHQNGKDIAAETLVK 198
 Db 185 KY---KNAKVPTLDE-----ILNRYGNANNTIETKSPDVTFGMKOL---LDTLD 229

Query 199 KGY---DKKTDWVYLOQFEDENELKRKTELLPQMGDLKLVQLIAYTDWKETOEKDPK 254
 Db 230 KHDLLTOKSLKHGHYMIQSFGSSLEKVKH----HMNANIPLTRIMNKPEFLKRNQDLIK 284
 Qy 255 GYWNTINYDWMFKPGAMAEVVKYADGVGPQWMLVNKEESKPDNIVYTPLVKELAQYNVE 314
 Db 285 -----NIKSYAIGVGFETYDL-----NIKNTRHLKNU---GFL 314
 Qy 315 VHPTVYTVRKDALPEFTDYNQMDALLNKGATGVFTDFPD 354
 Db 315 VHPTVYTVRKDALPEFTDYNQMDALLNKGATGVFTDFPD 343

Search completed: June 22, 2005, 11:42:18
 Job time : 45 secs

Qy 177 KAPWFHQNGKDIAAETLKVKY---GYDKKTDMVYLOQFDENELKRKTELLPQMGIM 232
 Db 178 KSPDVY---PMEMEQLASLRKKHLLNNNLKNGHMIQSFSDESLKTHRO----- 226

1 protein - protein search, using SW model									
Copyright (c) 1993 - 2005 Compugen Ltd.									
Sequence: 1 MDPTTLALSLAAAGVLAGCS.....YKEYDNKRIDSTSGHHHHHH 464									
Scoring table: BLOSUM62									
Gapext 0.5									
Searched: 1714042 seqs, 383979560 residues									
Total number of hits satisfying chosen parameters: 1714042									
Minimum DB seq length: 0									
Maximum DB seq length: 2000000000									
Last-processing: Minimum Match 0 %									
Maximum Match 10 %									
Listing First 45 summaries									
Database :									
Published Applications AA:*									
1:	/cgn2_6/_ptodata/2/_pubcaa/_US07_ PUBCOMB.pep.*								
2:	/cgn2_6/_ptodata/2/_pubcaa/_PCT_NEW_PUB.pep.*								
3:	/cgn2_6/_ptodata/2/_pubcaa/_US06_NEW_PUB.pep.*								
4:	/cgn2_6/_ptodata/2/_pubcaa/_US06_PUBCOMB.pep.*								
5:	/cgn2_6/_ptodata/2/_pubcaa/_US07_NEW_PUB.pep.*								
6:	/cgn2_6/_ptodata/2/_pubcaa/_PCTUS_PUBCOMB.pep.*								
7:	/cgn2_6/_ptodata/2/_pubcaa/_US08_NEW_PUB.pep.*								
8:	/cgn2_6/_ptodata/2/_pubcaa/_US08_PUBCOMB.pep.*								
9:	/cgn2_6/_ptodata/2/_pubcaa/_US09_ PUBLCOMB.pop.*								
10:	/cgn2_6/_ptodata/2/_pubcaa/_US09C_PUBLCOMB.pop.*								
11:	/cgn2_6/_ptodata/2/_pubcaa/_US09N_PUBLCOMB.pop.*								
12:	/cgn2_6/_ptodata/2/_pubcaa/_US10_N_PUBLCOMB.pop.*								
13:	/cgn2_6/_ptodata/2/_pubcaa/_US10A_PUBLCOMB.pop.*								
14:	/cgn2_6/_ptodata/2/_pubcaa/_US10B_PUBLCOMB.pop.*								
15:	/cgn2_6/_ptodata/2/_pubcaa/_US10C_PUBLCOMB.pop.*								
16:	/cgn2_6/_ptodata/2/_pubcaa/_US10D_PUBLCOMB.pop.*								
17:	/cgn2_6/_ptodata/2/_pubcaa/_US10E_PUBLCOMB.pop.*								
18:	/cgn2_6/_ptodata/2/_pubcaa/_US10_N_PUBLCOMB.pop.*								
19:	/cgn2_6/_ptodata/2/_pubcaa/_US11A_PUBLCOMB.pop.*								
20:	/cgn2_6/_ptodata/2/_pubcaa/_US11A_NEW_PUB.pep.*								
21:	/cgn2_6/_ptodata/2/_pubcaa/_US60_NEW_PUB.pep.*								
22:	/cgn2_6/_ptodata/2/_pubcaa/_US60_PUBLCOMB.pop.*								
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.									
SUMMARIES									
built	Query No.	Score	Match	Length	DB ID	Description	Length	358;	
1	1251.5	50.5	358	14	US-10-047-855-12	Sequence 12, Appl	50.5%;	Score 1251.5;	DB 14;
2	661.5	26.7	451	16	US-10-073-127-1113	Sequence 11,3, Appl	Best Local Similarity 64.1%;	Pred. No. 1.7e-93;	
3	651	26.3	144	17	US-10-687-060-7	Sequence 7, Appl	Matches 229;	Conservative 56;	Mismatches 65;
4	606	24.5	324	14	US-10-023-013-15	Sequence 15, Appl	Indels 7;	Gaps 2	
5	606	24.5	324	17	US-10-687-060-15	Sequence 15, Appl			
6	606	24.5	324	17	US-10-485-048-15	Sequence 15, Appl			
7	597	24.1	326	14	US-10-203-013-19	Sequence 19, Appl			
8	597	24.1	326	17	US-10-687-060-19	Sequence 19, Appl			
9	597	24.1	326	17	US-10-485-048-19	Sequence 19, Appl			
10	594.5	24.0	411	14	US-10-203-013-17	Sequence 17, Appl			
11	594.5	24.0	411	17	US-10-687-060-17	Sequence 17, Appl			
SUMMARIES									
built	Query No.	Score	Match	Length	DB ID	Description	Length	358;	
1	1251.5	50.5	358	14	US-10-047-855-12	Sequence 12, Appl	50.5%;	Score 1251.5;	DB 14;
2	661.5	26.7	451	16	US-10-073-127-1113	Sequence 11,3, Appl	Best Local Similarity 64.1%;	Pred. No. 1.7e-93;	
3	651	26.3	144	17	US-10-687-060-7	Sequence 7, Appl	Matches 229;	Conservative 56;	Mismatches 65;
4	606	24.5	324	14	US-10-023-013-15	Sequence 15, Appl	Indels 7;	Gaps 2	
5	606	24.5	324	17	US-10-687-060-15	Sequence 15, Appl			
6	606	24.5	324	17	US-10-485-048-15	Sequence 15, Appl			
7	597	24.1	326	14	US-10-203-013-19	Sequence 19, Appl			
8	597	24.1	326	17	US-10-687-060-19	Sequence 19, Appl			
9	597	24.1	326	17	US-10-485-048-19	Sequence 19, Appl			
10	594.5	24.0	411	14	US-10-203-013-17	Sequence 17, Appl			
11	594.5	24.0	411	17	US-10-687-060-17	Sequence 17, Appl			
SUMMARIES									
built	Query No.	Score	Match	Length	DB ID	Description	Length	358;	
1	1251.5	50.5	358	14	US-10-047-855-12	Sequence 12, Appl	50.5%;	Score 1251.5;	DB 14;
2	661.5	26.7	451	16	US-10-073-127-1113	Sequence 11,3, Appl	Best Local Similarity 64.1%;	Pred. No. 1.7e-93;	
3	651	26.3	144	17	US-10-687-060-7	Sequence 7, Appl	Matches 229;	Conservative 56;	Mismatches 65;
4	606	24.5	324	14	US-10-023-013-15	Sequence 15, Appl	Indels 7;	Gaps 2	
5	606	24.5	324	17	US-10-687-060-15	Sequence 15, Appl			
6	606	24.5	324	17	US-10-485-048-15	Sequence 15, Appl			
7	597	24.1	326	14	US-10-203-013-19	Sequence 19, Appl			
8	597	24.1	326	17	US-10-687-060-19	Sequence 19, Appl			
9	597	24.1	326	17	US-10-485-048-19	Sequence 19, Appl			
10	594.5	24.0	411	14	US-10-203-013-17	Sequence 17, Appl			
11	594.5	24.0	411	17	US-10-687-060-17	Sequence 17, Appl			
SUMMARIES									
built	Query No.	Score	Match	Length	DB ID	Description	Length	358;	
1	1251.5	50.5	358	14	US-10-047-855-12	Sequence 12, Appl	50.5%;	Score 1251.5;	DB 14;
2	661.5	26.7	451	16	US-10-073-127-1113	Sequence 11,3, Appl	Best Local Similarity 64.1%;	Pred. No. 1.7e-93;	
3	651	26.3	144	17	US-10-687-060-7	Sequence 7, Appl	Matches 229;	Conservative 56;	Mismatches 65;
4	606	24.5	324	14	US-10-023-013-15	Sequence 15, Appl	Indels 7;	Gaps 2	
5	606	24.5	324	17	US-10-687-060-15	Sequence 15, Appl			
6	606	24.5	324	17	US-10-485-048-15	Sequence 15, Appl			
7	597	24.1	326	14	US-10-203-013-19	Sequence 19, Appl			
8	597	24.1	326	17	US-10-687-060-19	Sequence 19, Appl			
9	597	24.1	326	17	US-10-485-048-19	Sequence 19, Appl			
10	594.5	24.0	411	14	US-10-203-013-17	Sequence 17, Appl			
11	594.5	24.0	411	17	US-10-687-060-17	Sequence 17, Appl			
SUMMARIES									
built	Query No.	Score	Match	Length	DB ID	Description	Length	358;	
1	1251.5	50.5	358	14	US-10-047-855-12	Sequence 12, Appl	50.5%;	Score 1251.5;	DB 14;
2	661.5	26.7	451	16	US-10-073-127-1113	Sequence 11,3, Appl	Best Local Similarity 64.1%;	Pred. No. 1.7e-93;	
3	651	26.3	144	17	US-10-687-060-7	Sequence 7, Appl	Matches 229;	Conservative 56;	Mismatches 65;
4	606	24.5	324	14	US-10-023-013-15	Sequence 15, Appl	Indels 7;	Gaps 2	
5	606	24.5	324	17	US-10-687-060-15	Sequence 15, Appl			
6	606	24.5	324	17	US-10-485-048-15	Sequence 15, Appl			
7	597	24.1	326	14	US-10-203-013-19	Sequence 19, Appl			
8	597	24.1	326	17	US-10-687-060-19	Sequence 19, Appl			
9	597	24.1	326	17	US-10-485-048-19	Sequence 19, Appl			
10	594.5	24.0	411	14	US-10-203-013-17	Sequence 17, Appl			
11	594.5	24.0	411	17	US-10-687-060-17	Sequence 17, Appl			
SUMMARIES									
built	Query No.	Score	Match	Length	DB ID	Description	Length	358;	
1	1251.5	50.5	358	14	US-10-047-855-12	Sequence 12, Appl	50.5%;	Score 1251.5;	DB 14;
2	661.5	26.7	451	16	US-10-073-127-1113	Sequence 11,3, Appl	Best Local Similarity 64.1%;	Pred. No. 1.7e-93;	
3	651	26.3	144	17	US-10-687-060-7	Sequence 7, Appl	Matches 229;	Conservative 56;	Mismatches 65;
4	606	24.5	324	14	US-10-023-013-15	Sequence 15, Appl	Indels 7;	Gaps 2	
5	606	24.5	324	17	US-10-687-060-15	Sequence 15, Appl			
6	606	24.5	324	17	US-10-485-048-15	Sequence 15, Appl			
7	597	24.1	326	14	US-10-203-013-19	Sequence 19, Appl			
8	597	24.1	326	17	US-10-687-060-19	Sequence 19, Appl			
9	597	24.1	326	17	US-10-485-048-19	Sequence 19, Appl			
10	594.5	24.0	411	14	US-10-203-013-17	Sequence 17, Appl			
11	594.5	24.0	411	17	US-10-687-060-17	Sequence 17, Appl			
SUMMARIES									
built	Query No.	Score	Match	Length	DB ID	Description	Length	358;	
1	1251.5	50.5	358	14	US-10-047-855-12	Sequence 12, Appl	50.5%;	Score 1251.5;	DB 14;
2	661.5	26.7	451	16	US-10-073-127-1113	Sequence 11,3, Appl	Best Local Similarity 64.1%;	Pred. No. 1.7e-93;	
3	651	26.3	144	17	US-10-687-060-7	Sequence 7, Appl	Matches 229;	Conservative 56;	Mismatches 65;
4	606	24.5	324	14	US-10-023-013-15	Sequence 15, Appl	Indels 7;	Gaps 2	
5	606	24.5	324	17	US-10-687-060-15	Sequence 15, Appl			
6	606	24.5	324	17	US-10-485-048-15	Sequence 15, Appl			
7	597	24.1	326	14	US-10-203-013-19	Sequence 19, Appl			
8	597	24.1	326	17	US-10-687-060-19	Sequence 19, Appl			
9	597	24.1	326	17	US-10-485-048-19	Sequence 19, Appl			
10	594.5	24.0	411	14	US-10-203-013-17	Sequence 17, Appl			
11	594.5	24.0	411	17	US-10-687-060-17	Sequence 17, Appl			
SUMMARIES									
built	Query No.	Score	Match	Length	DB ID	Description	Length	358;	
1	1251.5	50.5	358	14	US-10-047-855-12	Sequence 12, Appl	50.5%;	Score 1251.5;	DB 14;
2	661.5	26.7	451	16	US-10-073-127-1113	Sequence 11,3, Appl	Best Local Similarity 64.1%;	Pred. No. 1.7e-93;	
3	651	26.3	144	17	US-10-687-060-7	Sequence 7, Appl	Matches 229;	Conservative 56;	Mismatches 65;
4	606	24.5	324	14	US-10-023-013-15	Sequence 15, Appl	Indels 7;	Gaps 2	
5	606	24.5	324	17	US-10-687-060-				

Db 120 GFDIENCKKVQYTPGRPMGSDFRVTHFEEIEFYQGLNHTSGRNGITGYPEIKAPWFFH 179
 Qy 184 ONGKDAETEKLKVKKYDVKTDMYLQTFDFNELEKTRKTELLPOMGMDLKLVOLIAYT 243
 Db 180 QEGKDIAAKTLEVKKYYGTGDKVYQLQCTADELKRNKEELPEPMGMLNLVOLIAYT 239
 Qy 244 DWKETOEDPKPCQYWNWNYDNMFKPGAMAEMVYKVAADGVGPWYMLVNKEERSKPDNIVYTP 303
 Db 240 DNNETQKQPDSEWNTNDMFKPGAMKQVAEYGIGDPYHMI-EBTSGPNKLRG 298
 Qy 304 LYKELAQNYEVHYPYTKRDLPEFTDWNQMYDALINKSGATGVFTDFDTGVYFL 360
 Db 299 MVQDAAQCNKLVYHPYTVRSKDLPEFTPDVNQLYDALYNKACVNGLFTDFDKAVKFL 355

RESULT 2 US-10-473-127-1113

; Sequence 1113; Application US/10473127
 ; GENERAL INFORMATION:
 ; APPLICANT: Zycos Inc.
 ; TITLE OF INVENTION: TRANSLATIONAL PROFILING

; FILE REFERENCE: 08191-026W01
 ; CURRENT APPLICATION NUMBER: US/10/473,127
 ; CURRENT FILING DATE: 2003-09-26
 ; PRIOR APPLICATION NUMBER: 60/279,495
 ; PRIOR FILING DATE: 2001-03-28
 ; PRIOR APPLICATION NUMBER: 60/292,544
 ; PRIOR FILING DATE: 2001-05-21
 ; PRIOR APPLICATION NUMBER: 60/310,801
 ; PRIOR FILING DATE: 2001-08-08
 ; PRIOR APPLICATION NUMBER: 60/326,370
 ; PRIOR FILING DATE: 2001-10-01
 ; PRIOR APPLICATION NUMBER: 60/335,780
 ; PRIOR FILING DATE: 2001-12-04
 ; PRIOR APPLICATION NUMBER: 60/358,985
 ; PRIOR FILING DATE: 2002-02-20
 ; NUMBER OF SEQ ID NOS: 2041
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 1113
 ; LENGTH: 451
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens

US-10-473-127-1113

Query Match 26.7%; Score 661.5; DB 16; Length 451;
 Best Local Similarity 37.5%; Pred. No. 3..3e-45;
 Matches 194; Conservative 51; Mismatches 152; Indels 121; Gaps 18;
 Qy 1 MDPKTIALSLIAAGVLAGCSSSHSNMANANTOMKSDKLIIIAHGRASGYLPHTLESKALAPA 60
 Db 1 MDPKTIALSLIAAGVLAGCSSSHSNMANANTOMKSDKLIIIAHGRASGYLPHTLESKALAPA 60
 Qy 61 QOADYLEQDIAMTKDGRLVVHDIFLDGLTDVAKKFPHHRKGRRYVVDFTLKEIQSLE 120
 Db 61 QOADYLEQDIAMTKDGRLVVHDIFLDGLTDVAKKFPHHRKGRRYVVDFTLKEIQSLE 120
 Qy 61 QOADYLEQDIAMTKDGRLVVHDIFLDGLTDVAKKFPHRKDGRRYVVDFTLKEIQSLE 120
 Db 61 QOADYLEQDIAMTKDGRLVVHDIFLDGLTDVAKKFPHRKDGRRYVVDFTLKEIQSLE 120
 Qy 121 MTENFETKDGKAQYYPNRFLWKSHRHTTFDETEFLIGLERKVGKVIYPEIKAPW 180
 Db 121 MTENFETMDLBORSQ---HCKPBE---GLE ARGEALGLV-GAQAPA 159
 Qy 181 FHHONGKGDIAETLKVKKYDVKTDMYLQTFDFNELEKTRKTELLPQMGMDL-KLVQL 239
 Db 160 TEEQEAASSSTLVEY----TLGEVPAAESPPDPPSPQGASSLPFT 201
 Qy 240 IAYTKDWTKEQ---KDPKGYWWNNYDNMFPKG---AMAEVV----KY----AD 279
 Db 202 MNYPWPSQSYDDSSNQEEGPSTPQDLESEQQAALSRSKVAELVHFLLCYRAREPVTAK 261
 Qy 280 GVGP---GWMLVNKEESKPDNIVTPLKELAQNVEWHPYTVRKDALPEFFDVQMY 336
 Db 262 MLGSVVGWNQYFPVPIFSKASSSLQLVFGEIELMEVPIGHLYI-----FATCLGLSY 313

RESULT 3 US-10-687-060-7

; Sequence 7; Application US/10687060
 ; Publication No. US20050033022A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bruck, Claudine
 ; APPLICANT: Godart, Stéphane Andre Georges
 ; APPLICANT: Marc-Hanck, Martine
 ; TITLE OF INVENTION: Fusion Proteins Comprising HIV-1 Tat
 ; FILE REFERENCE: B45110
 ; CURRENT APPLICATION NUMBER: US/10/687,060
 ; CURRENT FILING DATE: 2003-10-16
 ; PRIOR APPLICATION NUMBER: US/09/509,239
 ; PRIOR FILING DATE: 2000-03-23
 ; PRIOR APPLICATION NUMBER: PCT/EP98/06040
 ; PRIOR FILING DATE: 1998-09-17
 ; PRIOR APPLICATION NUMBER: GB 9720585.0
 ; PRIOR FILING DATE: 1997-09-26
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: FastSSQ for Windows Version 3.0
 ; SEQ ID NO: 7
 ; LENGTH: 144
 ; TYPE: PRT
 ; ORGANISM: Pichia pastoris
 ; SEQ ID NO: 7
 ; LENGTH: 144;
 ; Query Match 26.3%; Score 651; DB 17; Length 144;
 ; Best Local Similarity 100.0%; Prod. No. 4.6e-45;
 ; Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 1 MDPKTIALSLIAAGVLAGCSSSHSNMANANTOMKSDKLIIIAHGRASGYLPHTLESKALAPA 60
 Db 1 MDPKTIALSLIAAGVLAGCSSSHSNMANANTOMKSDKLIIIAHGRASGYLPHTLESKALAPA 60
 Qy 61 QOADYLEQDIAMTKDGRLVVHDIFLDGLTDVAKKFPHRKDGRRYVVDFTLKEIQSLE 120
 Db 61 QOADYLEQDIAMTKDGRLVVHDIFLDGLTDVAKKFPHRKDGRRYVVDFTLKEIQSLE 120
 Qy 121 MTENFET 127
 Db 121 MTENFET 127
 ; Sequence 14; Application US/10203013
 ; Publication No. US20030158134A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SmithKline Beecham Biologicals S.A.
 ; TITLE OF INVENTION: No. US20030158134A1
 ; FILE REFERENCE: B45209
 ; CURRENT APPLICATION NUMBER: US/10/203,013
 ; CURRENT FILING DATE: 2002-07-31
 ; PRIOR APPLICATION NUMBER: GB 00022004
 ; PRIOR FILING DATE: 2000-01-31
 ; PRIOR APPLICATION NUMBER: GB 0009336.9
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: GB 0013806.5

PRIOR FILING DATE: 2000-06-06
PRIOR APPLICATION NUMBER: PCT/EP00/05998
PRIOR FILING DATE: 2000-06-28
NUMBER OF SEQ ID NOS: 31
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO: 15
LENGTH: 324
TYPE: PRT
ORGANISM: Homo sapiens
-10-203-013-15

Query: March 24 5%; Score: 606;
 Beat Local Similarity 37.3%; Pred. No.: Mismatches 168; Conservative 34; Mismatches 19

19	CSSHHSSNNNTQMSDKIIIAHRSAGSYLP	1	CSSHHSSNNNTQMSDKIIIAHRSAGSYLP
79	VYTHDHFIDLGTDVAKFPHRKDGRYYV	79	VYTHDHFIDLGTDVAKFPHRKDGRYYV
61	VYTHDHFIDLGTDVAKFPHRKDGRYYV	61	VYTHDHFIDLGTDVAKFPHRKDGRYYV
139	RFLWKSFRHTPFEDELEFICQLEKSTGK	139	RFLWKSFRHTPFEDELEFICQLEKSTGK
114	-W-----SKSS	114	-W-----SKSS
199	KYGDKTCIDMVLQTFDFNBLKRKIKTELLP	199	KYGDKTCIDMVLQTFDFNBLKRKIKTELLP
148	KRCAITTSNTAATN-:::	148	KRCAITTSNTAATN-:::
259	NNYNDWMFKPGAMAEVVYKYADGVGPWTYML	259	NNYNDWMFKPGAMAEVVYKYADGVGPWTYML
180	T-----PQVPLRPMTYKAAYDLSHPL-:-	180	T-----PQVPLRPMTYKAAYDLSHPL-:-
319	TYRKDALBEPFDVNMQMDALLNSGATQV	319	TYRKDALBEPFDVNMQMDALLNSGATQV
215	--RQDIDLWLYHTQ-----GY	215	--RQDIDLWLYHTQ-----GY
375	LVRSDYKLKNQSSNNSNSTLNLGBHHRARA	375	LVRSDYKLKNQSSNNSNSTLNLGBHHRARA
256	PYPEPD-KVEEPAKGENTSLLPHPSLHL--G	256	PYPEPD-KVEEPAKGENTSLLPHPSLHL--G
435	RHTARTSRSDYKPYDNKRIDSSTSGHHHHHH	435	RHTARTSRSDYKPYDNKRIDSSTSGHHHHHH
301	RHVARELHPPEY-FKN----CTSGHAAHHHH	301	RHVARELHPPEY-FKN----CTSGHAAHHHH

SU(NT 5
-10-687-060-15
Sequence 15, Application US/10687060
Publication No. US20050033022A1
GENERAL INFORMATION:
APPLICANT: Bruck, Claudine
APPLICANT: Godart, Stephane Andre George
APPLICANT: Marc-Hand, Martine
TITLE OF INVENTION: Fusion Proteins Comprising
Title of Invention: and/or Nef Proteins
FILE REFERENCE: B45110
CURRENT APPLICATION NUMBER: US/10687,060
CURRENT FILING DATE: 2003-10-16
PRIOR APPLICATION NUMBER: US/09/509,239
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/EP98/06040
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: GB 9720585.0
PRIOR FILING DATE: 1997-09-26
NUMBER OF SEQ ID NOS: 27
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO: 15
LENGTH: 324
TYPE: PRT
ORGANISM: Pichia pastoris

US-10-587-060-15	Query Match	Best Local Sim	Matches	168;
	Qy	19 CS- 		
	Db	1 CS- 		
	Qy	79 VV- 		
	Db	61 VV- 		
	Qy	139 RF- 		
	Db	114 -- 		
	Qy	199 KYC- : 		
	Db	148 KHK- 		
	Qy	259 NYM- 		
	Db	180 T- 		
	Qy	319 TVD- 		
	Db	215 -- 		
	Qy	375 LVI- 		
	Db	256 PVI- 		
	Qy	435 RHI- 		
	Db	301 HRY- 		
RESULT 6				
US-10-485-048-15				
		Publication No		
		GENERAL INFORMATION		
		APPLICANT: ERIC		
		APPLICANT: V		
		TITLE OF INVENTION		
		TITLE OF INVENTION		
		FILE REFERENCE		
		CURRENT APPLICATION		
		CURRENT FILING		
		PRIOR APPLICATION		
		PRIOR FILING		
		PRIOR APPLICATION		
		PRIOR FILING		
		NUMBER OF SEQ		
		SOFTWARE: Fast		
		SEQ ID NO 15		
		LENGTH: 324		
		TYPE: PRT		
		ORGANISM: Hor		
US-10-485-048-15				
Query Match				
Best Local Sim				
Matches 168;				

Qy 79 VVIIDHFLGLTDVAKKPFPHRKDRYVVIDFTLKEIQSLEMTEMFETKGKQAVVYN 138
 Db 61 VVIIDHFLGLTDVAKKPFPHRKDRYVVIDFTLKEIQSLEMTEMFETGGK----- 113

Db 139 RFLPKSHFRHTFEDEIEFTGLEKSTGKVGTYPEIKAPWFHQNQGKDIAETLKVLK 198
 Db 114 ---W-----SKSSVVG-WPTVRMRRAPEAAGVGAASRDLK 147

Qy 199 KYGYDKKTMVLTQDFTNELLKRKITELLPQMGMILKLKVQJAYTDWKETOEKDPKGWY 258
 Db 148 KHGATTSNTAATN-----AACAWLQAQEEEVGFV 179

Qy 259 NYNDNNPKGAMAEVVKYADGVPMVNKEESKPDNIVTPLYKELAQYNNVEHPY 318
 Db 180 T-----PVQLRMTYKAVDLSHF-----PLTFGWCYKLV 214

Qy 319 TVRKDALPEFTDVNQMYDALLNKSGATGVFTDF-----PTGVEFLKGTKSMDGGKAGVA 374
 Db 215 -RQDLDLWYHTQ-----GFPDWQNYTPGPGRVY-----PLTFGWCYKLV 255

RESULT 8
 US-10-687-060-19 ; Sequence 19, Application US/10587060
 Db 256 PVEPD-KVEEANKGENTSLHPVSLH--GMDPEREVLBWRFDRL-----AP 300 ; GENERAL INFORMATION:
 ; APPLICANT: Bruck, Claudine
 ; APPLICANT: Godart, Stephane Andre Georges
 ; APPLICANT: Marc-Hard, Martine
 ; TITLE OF INVENTION: Fusion Proteins Comprising HIV-1 TAT
 ; FILE REFERENCE: B45110
 ; CURRENT APPLICATION NUMBER: US/10/687,060
 ; PRIOR APPLICATION NUMBER: US/09/509,239
 ; PRIOR FILING DATE: 2000-03-23
 ; PRIOR APPLICATION NUMBER: PCT/EP98/06040
 ; PRIOR FILING DATE: 1998-09-17
 ; PRIOR APPLICATION NUMBER: GB 9720585.0
 ; PRIOR FILING DATE: 1997-09-26
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO 19
 ; LENGTH: 326
 ; TYPE: PRT
 ; ORGANISM: Pichia pastoris
 US-10-687-060-19

Query Match 24.1%; Score 597; DB 17; Length 326;
 Best Local Similarity 37.2%; Pred. No. 3.8e-40;
 Matches 167; Conservative 34; Mismatches 118; Indels 130; Gaps 15;

Db 20 SSHSSMNANTQMSKIIIAHRGASGYLPEHTLESKALAFQQADYLEDQDLAMTKDGLV 79
 Db 4 SSHSSMNANTQMSKIIIAHRGASGYLPEHTLESKALAFQQADYLEDQDLAMTKDGLV 63

Qy 80 VTHDHFGLDGTDVAKKPFPHRKDRYVVIDFTLKEIQSLEMTEMFETKGKQAVQYVNR 139
 Db 64 VTHDHFGLDGTDVAKKPFPHRKDRYVVIDFTLKEIQSLEMTEMFETGGK----- 115

Qy 140 FPLWKSHFRHTFEDEIEFTGLEKSTGKVGIVYPIKAPWFHQNQGKDIAETLKVLK 199
 Db 116 ---W-----SKSSVVG-WPTVRMRRAPEAAGVGAASRDLK 150

Qy 200 YGDKKTDMVLYLQFTDFNELKRIKTELLPQMGMILKLVOLIAYTDWKETOEKDPKGWYN 259
 Db 151 HGATTSNTAATN-----AACAWLQAQEEEVGFV 182

Qy 260 YNDNMFKGAMAEVVKYADGVPMVNKEESKPDNIVTPLYKELAQYNNVEHPY 319
 Db 183 -----PVQLRMTYKAVDLSHF-----PLTFGWCYKLV 216

Db 116 ---W-----SKSSVVG-WPTVRMRRAPEAAGVGAASRDLK 150
 Qy 200 YGDKKTDMVLYLQFTDFNELKRIKTELLPQMGMILKLVOLIAYTDWKETOEKDPKGWYN 259
 Db 217 -RQDLDLWYHTQ-----GFPDWQNYTPGPGRVY-----PLTFGWCYKLV 258

Qy 376 VRSVDYKLYNKNSSSNSTLKNLGEHHRARAMDGGKAGVALVRSVDYKFYE DANGTRDHKKGR 435
Db 259 VEPD-KVEANKGENTSLHPVSLH---GMDDPEREVLWRFDSRL-----AFH 303

Qy 436 HTPARTSRSDYKFDNKRDLSTSGHHHHH 464
Db 304 HVARELHPEY--FKN---CTSGHHHHHH 326

RESULT 9
US-10-485-048-19
; Sequence 19, Application US/10485048
; Publication No. US2005005865781
; GENERAL INFORMATION
; APPLICANT: ERTL, Peter Franz
; TITE: John Phillip
; APPLICANT: VAN WEY, Catherine Anne
; APPLICANT: VOSS, Gerald
; TITLE OF INVENTION: Vaccine Comprising GP120 and NEF and/or
; FILE REFERENCE: B45279
; CURRENT APPLICATION NUMBER: US/10/485,048
; PRIOR FILING DATE: 2004-01-26
; PRIOR APPLICATION NUMBER: PCT/EP02/08343
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: GB 0118367.2
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 19
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Homo sapien
; US-10-485-048-19

Query Match 24.1%; Score 597; DB 17; Length 326;
Best Local Similarity 37.2%; Pred. No. 3. 8e-10;
Matches 167; Conservative 34; Mismatches 118; Indels 130; Gaps 15;

Qy 20 SSSHSSNNANTQMSDKIIIAHRSAGSGLPEHTLESKALAAFAQAQADYLEODLAMTKDGLR 79
Db 4 SSSHSSNNANTQMSDKIIIAHRSAGSGLPEHTLESKALAAFAQAQADYLEODLAMTKDGLR 63

Qy 80 VTHDHFLDGLTDVAKKEPKPHRKDKGYYVIFTLKRISLEMTENPETDKQKAQVYPR 139
Db 64 VTHDHFLDGLTDVAKKFPHRKDKGYYVIFTLKRISLEMTENPETMGSK-----115

Qy 140 FPLWKSHPRHTFDETEFIQGLEKSTGKCYIPIKA PMFHONGKIDAAETLKVLLK 199
Db 116 ---W-----SKSWSVG-WPTVTRMRRAEPADGVGAARDLKE 150

Qy 200 YGYDKKTDMMYLQTFDFNELKRITBELLPKQDKPGYWN 259
Db 151 HGAITSNTAATN-----AACAWLEAQEEBEGFPPT 182

Qy 260 YNYDMFKPGAMAEVVKYADCGPGWMLVNKEESKPDKNIVYTPLVKELQYNVEVHPYT 319
Db 183 -----POVPLRPMTYKAADVLSHFL---KEKGSLLEGJHSQR-----216

Qy 320 VRKDALPEFFPDVNQMYDALLNKSAGATGVFTDF---PDTGVEFLKGIGSMDDGKAGVAL 375
Db 217 -RQDILWLWYHTQ-----GYFPDWQNYTPGPGVRY---PLTFGWCYKLVP 258

Qy 376 VRSVDYKLYNKNSSSNSTLKNLGEHHRARAMDGGKAGVALVRSVDYKFYE DANGTRDHKKGR 435
Db 259 VEPD-KVEANKGENTSLHPVSLH---GMDDPEREVLWRFDSRL-----AFH 303

Qy 436 HTPARTSRSDYKFDNKRDLSTSGHHHHH 464
Db 304 HVARELHPEY--FKN---CTSGHHHHHH 326

RESULT 10
US-10-203-013-17
; Sequence 17, Application US/10203013
; Publication No. US20030158134A1
; GENERAL INFORMATION
; APPLICANT: SmithKline Beecham Biologicals S.A.
; TITLE OF INVENTION: No. US20030158134A1el Use
; FILE REFERENCE: B45209
; CURRENT APPLICATION NUMBER: US/10/203,013
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: GB 0002200-4
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: GB 0009336-9
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: GB 0013805-5
; PRIOR FILING DATE: 2000-05-06
; PRIOR APPLICATION NUMBER: PCT/EP00/05998
; PRIOR FILING DATE: 2000-05-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO: 17
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-203-013-17

Query Match 24.0%; Score 594.5; DB 14; Length 411;
Best Local Similarity 35.6%; Pred. No. 8.4e-40;
Matches 178; Conservative 45; Mismatches 134; Indels 143; Gaps 19;

Qy 19 CSSHSSNNANTQMSDKIIIAHRSAGSGLPEHTLESKALAAFAQAQADYLEODLAMTKDGLR 78
Db 1 CSSHSSNNANTQMSDKIIIAHRSAGSGLPEHTLESKALAAFAQAQADYLEODLAMTKDGLR 60

Qy 79 VVTHDHFLDGLTDVAKKEPKPHRKDKGYYVIFTLKRISLEMTENPETDKQKAQVYPR 138
Db 61 VVTHDHFLDGLTDVAKKFPHRKDKGYYVIFTLKRISLEMTENPETMGK-----113

Qy 139 RFLPKSHPRHTFDETEFIQGLEKSTGKCYIPIKA PMFHONGKIDAAETLKVLLK 198
Db 114 ---W-----SKSSSVG-WPTVTRMRRAEPADGVGAARDLDE 147

Qy 199 KYGYDKKTD-----MVYLQTDFNELKRITBELLPKQDGMLK-LVOLJAYTDWK-----246
Db 148 KHGATTSNTAATNACAWLQAQBEBVGSPVTPVPLRPMTYKAADVLSHFLKGGLB 207

Qy 247 ---ETQEKDPKGTVW---NNYDMMMPKGAMAEVVKYADGVGP----GW-YMLVN- 290
Db 208 GLIHQRQRDILWYHTQGYFPFWQ-----NVTPGPGVTRYPLTFGWCYKLVPV 257

Qy 291 -----KEEKSKPDNTIVYTPLVKELA-----QVNVEVPTVTRKDALBFPTD 331
Db 258 EPDKKEEARNGEN---TSLLHPVSLHGMDDPEREWLEWRFDSRLAHPHYARELHPYPRN 314

Qy 332 VNQMDALLNKSAGATGVFTDFP-DTGVBLPKGIGKSMDDGKAGVALVRSVDYKLYNNNSSN 390
Db 315 C-----TSBPVDRPLBPMKHPSSQP-----KTACTN 340

Qy 391 STLNKLGEH----HRADAMDGSKAGVALVRSVDYKFYE DANGTRDHKKGRHTARTSRSD 444
Db 341 CYCKCCFCQCVCPFTKALGJSYGR-----KCRQRQRREPQGSATHOVSLSKQTSQSR 394

Qy 445 YKPYDNKRLDSTSIGHHHHH 464
Db 395 ---GBPTGPKETSGHHHHHH 411

RESULT 11
US-10-68-060-17
; Sequence 17, Application US/10203013
; Publication No. US20050033022A1
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine

APPLICANT: Godart, Stephane Andre Georges
 APPLICANT: Marc-Han, Martine
 TITLE OF INVENTION: Fusion Proteins Comprising HIV-1 TAT
 TITLE OF INVENTION: and/or Nef Proteins
 FILE REFERENCE: B45279
 CURRENT APPLICATION NUMBER: US/10/485,048
 CURRENT FILING DATE: 2004-01-26
 PRIORITY NUMBER: PCT/EP02/08343
 CURRENT APPLICATION NUMBER: US/10/687,060
 CURRENT FILING DATE: 2002-07-26
 PRIORITY NUMBER: GB 0118367.2
 PRIORITY FILING DATE: 2001-07-27
 PRIORITY FILING DATE: 2000-03-23
 PRIORITY FILING DATE: 1998-09-17
 PRIORITY APPLICATION NUMBER: GB 9720585.0
 PRIORITY FILING DATE: 1997-09-26
 NUMBER OF SEQ ID NOS: 98
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 17
 LENGTH: 411
 TYPE: PRT
 ORGANISM: Pichia pastoris
 US-10-687-06-17

Query Match 24.0%; Score 594.5; DB 17; Length 411;
 Best Local Similarity 35.6%; Pred. No. 8.4e-40;
 Matches 178; Conservative 45; Mismatches 134; Indels 143; Gaps 19;

Qy 19 CSSHSSNMANTQMSKDKIIIAHRSAGSYLPEHTLESKALAAQQADYLEQDLAMTKDGRLL 78
 Db 1 CSSHSSNMANTQMSKDKIIIAHRSAGSYLPEHTLESKALAAQQADYLEQDLAMTKDGRLL 60
 Qy 79 VVTHDHFLGILTDAVAKFPHRHKDGRYYTIDFTLKEIOSLEMENFETKDGKQAQVYPN 138
 Db 61 VVTHDHFLGILTDAVAKFPHRHKDGRYYTIDFTLKEIOSLEMENFETMGGK-----113
 Qy 139 RPLWKSHSRHTFEDEIPIQGLEKSTGKVG1YPEIKAPWPHHQNGDIAAEFLKVLK 198
 Db 114 ----W-----114
 Qy 199 KGYDKRTD-----MVYLQTEFDNEELKRKITLELPONGMDLK-LVQLJAYTDWK----246
 Db 148 KHGAISSNTAATNAACAWLEAQEEBEGFPVTQVPLRPMTYKAADVLSHELEKGGLLE 207
 Qy 247 ----ETQEKDPKGYWV----NNYNDMPKPGAMAEEVVKYADGVGP----GW-YMLVN-
 Db 208 GLHSQRQDILDLWYHTQGYFPDWQ-----NYTPGPGRYPLTFGWCYKLVPV 257
 Qy 291 ----KEESKPDNIVTPLYKELA-----QYNVEVHPYTVRKDALPEFFTD 331
 Db 258 EPDKVBEANKGEN--TSUJHPVSLHGMDDPEREYLEWREDSRLAFHAYARELPEYFQN 314
 Qy 332 VNQMYDALLNSKGATGVFTDPP-DTGVEFLKGKSMDGSKAGVALVRSYDLYKYNKSSSN 390
 Db 315 C-----TSEPVDPRLPEWKHPESSQP-----KTACTN 340
 Qy 391 STLNKLGEH-----HRADAMDGKTAGVALVRSYDLYKYNKSSSN 390
 Db 341 CYCKKCCPHICQVCPFTKALGISYGR-----KRRQRRRPQGSQTHQVSLSKOPTSQSR 394
 RESULT 13
 US-10-203-013-21
 Sequence 21, Application US/10203013
 Publication No. US20030158134A1
 GENERAL INFORMATION:
 APPLICANT: SmithKline Beecham Biologicals S.A.
 TITLE OF INVENTION: No. US20030158134A1 Use
 FILE REFERENCE: B45209
 CURRENT APPLICATION NUMBER: US/10/203,013
 CURRENT FILING DATE: 2002-07-31
 PRIORITY NUMBER: GB 00022004
 PRIORITY FILING DATE: 2000-01-31
 PRIORITY APPLICATION NUMBER: GB 0009336.9
 PRIORITY FILING DATE: 2000-04-14
 PRIORITY APPLICATION NUMBER: GB 0013806.5
 PRIORITY FILING DATE: 2000-06-06
 PRIORITY APPLICATION NUMBER: PCT/EP00/05998

RESULT 12
 US-10-485-048-17
 Sequence 17, Application US/10485048
 Publication No. US20050058657A1
 GENERAL INFORMATION:
 APPLICANT: ERTL, Peter Franz
 APPLICANT: TITE, John Phillip
 APPLICANT: VAN WELY, Catherine Anne
 APPLICANT: VOSS, Gerald
 TITLE OF INVENTION: Vaccine Comprising GP120 and NEF and/or
 TITLE OF INVENTION: TAT for the Immunisation Against HIV

PRIOR FILING DATE: 2000-06-28
 NUMBER OF SEQ ID NOS: 31
 SOFTWARE: FastSEQ for Windows Version 3.0
 SEQ ID NO: 21
 LENGTH: 413
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-203-013-21

Query Match Score 585.5; DB 14; Length 413;
 Best Local Similarity 35.5%; Pred. No. 4.6e-39;
 Matches 177; Conservative 45; Mismatches 134; Indels 143; Gaps 19;

Query 20 SSSHSSNANTOMKSDKIIIAHRGASGYLPENTLESKALAFAAQADYLEQDLAMTKDGRLV 79
 Db 4 SSSHSSNANTOMKSDKIIIAHRGASGYLPENTLESKALAFAAQADYLEQDLAMTKDGRLV 63

Query 20 SSSHSSNANTOMKSDKIIIAHRGASGYLPENTLESKALAFAAQADYLEQDLAMTKDGRLV 79
 Db 4 SSSHSSNANTOMKSDKIIIAHRGASGYLPENTLESKALAFAAQADYLEQDLAMTKDGRLV 63

Db 80 VHDFHFLDGLTDVAKCPFHRRKGDRYYVDFTLKEIQTSLLENTEPDKGQAOQVYPNR 139
 Query 80 VHDFHFLDGLTDVAKCPFHRRKGDRYYVDFTLKEIQTSLLENTEPDKGQAOQVYPNR 139
 Db 64 VHDFHFLDGLTDVAKCPFHRRKGDRYYVDFTLKEIQTSLLENTEPDMGK----- 115
 Query 200 YGYDKKTD-----MVLQLQTDPFDNLKRLKXTELLPQMGMDLK-LYOLIAYTDWK----- 246
 Db 116 ---W-----SKSISVG-WPTVRERMRRAEPADGVAASRDLEK 150

Query 140 FPLWKSHFRITHFEDIEFIQGLEKSTGKKGTYPEIKAPMFHHQNGKDAETLKVKK 199
 Db 116 ---W-----SKSISVG-WPTVRERMRRAEPADGVAASRDLEK 150

Query 200 YGYDKKTD-----MVLQLQTDPFDNLKRLKXTELLPQMGMDLK-LYOLIAYTDWK----- 246
 Db 151 HGAITSSNTAATNAACAWLAQEEBEVGFPTWPQPLRPMTYKAAVDLSHFLKERCGLEG 210

Query 247 ---ETOBKDPKGYW-----NNYNDMFKPGAMAETVVKYADGVGP-----GW-YMLVN-- 290
 Db 211 LIHSQRQDILDLWLYHTQGYFPDIO-----NYTPQPGVTPYRPLTFCWYKLVPVE 260

Query 291 ---KEESKPDNIVYTPVKELA-----QYNEVHPYTPVTKDALTSHPLKERCGLEG 210
 Db 261 PDKVEEANGEN---TSLLHPVSLHGMDDPEREYLEWRFDSSLRFAHVARELHPFKNC 317

Query 247 ---ETOBKDPKGYW-----NNYNDMFKPGAMAETVVKYADGVGP-----GW-YMLVN-- 290
 Db 333 NQMYDALLNKSGATGYPTDPP-DIGYBFLRGKIKSMDDGKGAGVALRSVDKLYNKSSNS 391

Query 291 ---KEESKPDNIVYTPVKELA-----QYNEVHPYTPVTKDALTSHPLKERCGLEG 210
 Db 318 -----TSEPVDPRLPEWKHPGSQP-----KTACTNC 343

Query 333 NQMYDALLNKSGATGYPTDPP-DIGYBFLRGKIKSMDDGKGAGVALRSVDKLYNKSSNS 391
 Db 392 TLKNLGEH-----HRARAMDGKGAGVALRSVDYKEYEDANGTRDKKGRTARTSRSDY 445

Query 392 TLKNLGEH-----HRARAMDGKGAGVALRSVDYKEYEDANGTRDKKGRTARTSRSDY 445
 Db 344 YCKKCCPHCQVCPIKALGISYGR-----KKRQRORRPPQGSQTHVSLSKQPTQSQR- 396

Query 446 KFYDNKRIDSSTSGHHHHHH 464
 Db 397 -GDPtgPKETSGHHHHHH 413

RESULT 155
 US-10-485-048-21
 Sequence 21, Application US/10485048
 Publication No. US2005058557A1
 GENERAL INFORMATION:
 APPLICANT: ERTL, Peter Franz
 APPLICANT: TITE, John Phillip
 APPLICANT: VAN WELY, Catherine Anne
 APPLICANT: VOSS, Gerald
 TITLE OF INVENTION: Vaccine Comprising GP120 and NEP and/or TAT for the Immunisation Against HIV
 FILE REFERENCE: B45279
 CURRENT APPLICATION NUMBER: US/10/485,048
 CURRENT FILING DATE: 2004-01-26
 PRIOR APPLICATION NUMBER: PCT/EP02/08343
 PRIOR FILING DATE: 2002-07-26
 PRIOR APPLICATION NUMBER: GB 01183367.2
 PRIOR PILING DATE: 2001-07-27
 NUMBER OF SEQ ID NOS: 98
 SEQ ID NO: 21
 LENGTH: 413
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-485-048-21

RESULT 14
 US-10-687-060-21
 Sequence 21, Application US/10687060
 Publication No. US050533022A1
 GENERAL INFORMATION:
 APPLICANT: Bruck, Claudine
 APPLICANT: Godart, Stephane Andre Georges
 APPLICANT: Marc-Han, Martine
 TITLE OF INVENTION: Fusion Proteins Comprising HIV-1 TAT
 TITLE OF INVENTION: and/or Nef Proteins
 FILE REFERENCE: B45110
 CURRENT APPLICATION NUMBER: US/10/687,060
 CURRENT FILING DATE: 2003-10-16
 PRIOR APPLICATION NUMBER: US/09/509,239
 PRIOR FILING DATE: 2000-03-23
 PRIOR APPLICATION NUMBER: PCT/EP08/06040
 PRIOR FILING DATE: 1998-09-17
 PRIOR APPLICATION NUMBER: GB 9720585.0
 PRIOR FILING DATE: 1997-09-26
 NUMBER OF SEQ ID NOS: 27
 SEQ ID NO: 21
 LENGTH: 413
 SOFTWARE: FastSEQ for Windows Version 3.0
 SOFTWARE: FastSEQ for Windows Version 3.0
 LENGTH: 413;

Matches	177;	Conservative	45;	Mismatches	134;	Indels	143;	Gaps	19;
Qy	20	SSHSSNMANTOMKSDKIIIAHRGASGYLPEHTLESKALAAFAQAQADYLEODLAMTKDGRIV	79						
	4	SSHSSNMANTOMKSDKIIIAHRGASGYLPEHTLESKALAAFAQAQADYLEODLAMTKDGRIV	63						
Db	80	VIHDHELDGLTDVAKEPFHHRKGYYVDFTLKEIISLEMTEFETDKQAOVYPNR	139						
Qy	64	VIHDHELDGLTDVAKEPFHHRKGYYVDFTLKEIISLEMTEFETDKQAOVYPNR	115						
Db	140	FPLWKSFRHTFEDIEFIGLEKSTGKKVGIYPEIKAPMFHHONGKODIAETTLKVLCK	199						
Qy	116	-----SKSUVG-WPTVRMRMRAADGVGAASRDLEK	150						
Db	116	-----							
Qy	200	YGYDKKTD-----NVYLOTFDENFLKRKTELLPQMCDLKL-NVQLAYTDWK	246						
Db	151	HGAITSNTAATNAACAWLAEQEEBVGFPPTPOVBLRNPETYKAADVLSHFLKEBGGLEG	210						
Qy	247	--ETQEKDPKGYWN---NNTNDMFKPGAMAEVVKYANGVGP-----GW-YMLVN-	290						
Db	211	LHSQPRQDILDLWLYTHTOGYFPDNO-----NYTRPGVRYPLTGFWCYKLVPVE	260						
Qy	291	----KBEISKPDNIVTPVKELA-----QYNEVEHYTYTVRKDALPFFTDV	332						
Db	261	PDKVEEANKGEN--TSLLHPVSLHGMDPPEREYLEWRFDSRLATHVARELHPBYFNC	317						
Qy	333	NOMYDALLNSKGATGVTDFP-DTGVBELKSIKSMDDGGKAGVALYRSDYKLYNKNSNS	391						
Db	318	-----TSEPVDPDRLEPWHPGSQB-----KFACTNC	343						
Qy	392	TLNULGEH----HRARAMDGKGAGVALYRSDYKFEDANGTRDHKKGRHTARTSRSDY	445						
Db	344	YCKKCCFHCOVCFITKALGSIYGR-----KKRRQRRLPFGSOVHSKOPPSQSR-	396						
Qy	446	KFYDNKRIDSSTSGHHHHHHH	464						
Db	397	-GDPTRGPKETSGHHHHHHH	413						

Search completed: June 22, 2005, 11:54:14
Job time : 161 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: June 22, 2005, 11:31:10 ; Search time 41 Seconds

(without alignments)
1088.892 Million cell updates/sec

Title: US-09-719-379A-S1

Perfect score: 2477

Sequence: 1 MDPTKLALSLLAAGVLAGGS.....YKFYDNKRIDSTSGHHHHH 464

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_79;*

1: Pir1;*

2: Pir2;*

3: Pir3;*

4: pir4;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	1920	glycerophosphodiesters	glycerophosphodiesters	77.5	364 2 A43576	glycerophosphodiesters
2	1916	glycerophosphodiesters	glycerophosphodiesters	77.4	364 2 S55932	glycerophosphodiesters
3	1914	glycerophosphodiesters	glycerophosphodiesters	77.3	364 2 S55931	glycerophosphodiesters
4	1910	glycerophosphodiesters	glycerophosphodiesters	77.1	364 2 S55934	glycerophosphodiesters
5	1902	glycerophosphodiesters	glycerophosphodiesters	76.8	364 2 S60486	glycerophosphodiesters
6	1885	glycerophosphodiesters	glycerophosphodiesters	76.1	364 2 S55933	glycerophosphodiesters
7	1885	glycerophosphodiesters	glycerophosphodiesters	76.1	364 2 S55936	glycerophosphodiesters
8	1261.5	glycerophosphodiesters	glycerophosphodiesters	50.9	358 2 F85863	glycerophosphodiesters
9	1261.5	glycerophosphodiesters	glycerophosphodiesters	50.9	358 2 S15245	glycerophosphodiesters
10	1251.5	glycerophosphodiesters	glycerophosphodiesters	50.5	358 2 S10791	glycerophosphodiesters
11	1233	glycerophosphodiesters	glycerophosphodiesters	49.8	356 2 AC0410	glycerophosphodiesters
12	1210	glycerophosphodiesters	glycerophosphodiesters	48.8	371 2 F71346	probable glyceroph
13	1074	glycerophosphodiesters	glycerophosphodiesters	43.4	356 2 F71346	glycerophosphodiesters
14	957	glycerophosphodiesters	glycerophosphodiesters	38.6	371 2 H82497	glycerophosphodiesters
15	425	glycerophosphodiesters	glycerophosphodiesters	17.2	1027 2 AC1841	glycerophosphodiesters
16	415.5	glycerophosphodiesters	glycerophosphodiesters	16.8	786 2 D75610	glycerophosphodiesters
17	383	glycerophosphodiesters	glycerophosphodiesters	15.5	383 2 S83603	glycerophosphodiesters
18	368	glycerophosphodiesters	glycerophosphodiesters	14.9	356 2 P87554	glycerophosphodiesters
19	354	glycerophosphodiesters	glycerophosphodiesters	14.3	389 2 AH1937	glycerophosphodiesters
20	353	glycerophosphodiesters	glycerophosphodiesters	14.3	372 2 T45528	glycerophosphodiesters
21	347.5	glycerophosphodiesters	glycerophosphodiesters	14.0	392 2 D956770	glycerophosphodiesters
22	321.5	glycerophosphodiesters	glycerophosphodiesters	13.0	293 2 I40418	glycerophosphodiesters
23	269.5	glycerophosphodiesters	glycerophosphodiesters	10.9	309 2 H89662	glycerophosphodiesters
24	262.5	glycerophosphodiesters	glycerophosphodiesters	10.6	284 2 DB4112	glycerophosphodiesters
25	246.5	glycerophosphodiesters	glycerophosphodiesters	10.0	375 2 A83352	glycerophosphodiesters
26	222	glycerophosphodiesters	glycerophosphodiesters	9.0	243 2 E69227	glycerophosphodiesters
27	181	glycerophosphodiesters	glycerophosphodiesters	7.3	247 2 E89956	glycerophosphodiesters
28	179.5	glycerophosphodiesters	glycerophosphodiesters	7.2	239 2 F69361	glycerophosphodiesters
29	172.5	glycerophosphodiesters	glycerophosphodiesters	7.0	583 2 AH1151	glycerophosphodiesters

RESULT 1

A43576

Glycerophosphodiester phosphodiesterase (EC 3.1.4.46) precursor - Haemophilus influenzae

N; Alternative names: immunoglobulin D-binding protein; protein D

C; Species: Haemophilus influenzae

A; Variety: isolate 772

C; Date: 22-Jan-1993 #sequence_change 22-Jan-1993 #text_change 08-Oct-1999

C; Accession: A43576

R; Janson, H.; Heden, L.O.; Grubb, A.; Ruan, M.; Forsgren, A.

A; Title: Protein D, an immunoglobulin D-binding protein of Haemophilus influenzae: clone

A; Reference number: A43576; PMID:9109948; PMID:1987023

A; Molecule type: DNA

A; Residues: 1-364 <JAN>

A; Cross-references: GB:M7487; NID:9148970; PIDN:AAA24998.1; PID:9148971

A; Experimental source: isolate 772

C; Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen

F; 1-18/Domain: signal sequence predicted <SIG>

F; 19-364/Product: glycerophosphodiester phosphodiesterase #status predicted <MAT>

Query Match 77.5%; Score 1920; DB 2; Length 364;

Best Local Similarity 99.5%; Pred. No. 1.4e-122; Matches 362; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MDPKTLALSLLAAGTLAGCSSHSMNANTOMKSDKIIIAHARGASGYLPEHTLESKALAF 60

Db 1 MQLKTQLALSLLAAGTLAGCSSHSMNANTOMKSDKIIIAHARGASGYLPEHTLESKALAF 60

Qy 61 QADYLQEDLAMTKDGRLVVTHDFLDGLTDVAKKEPHRKDKGRVYVIDFTLKETOSL 120

Db 61 QQADYLQEDLAMTKDGRLVVTHDFLDGLTDVAKKEPHRKDKGRVYVIDFTLKETOSL 120

Qy 121 MTENFETKDGKQAOQYPNRFPLWKSHFRIFTFEDEBFFIGLEKSTGKGVLYPEIKAPW 180

Db 121 MTENFETKDGKQAOQYPNRFPLWKSHFRIFTFEDEBFFIGLEKSTGKGVLYPEIKAPW 180

Qy 181 FHHQNGKDIAETLKVLLKKGYDVKTDMVYLQTDFNLKRKTKTELLPOMGMNDLKLVQLI 240

Db 181 FHHQNGKDIAETLKVLLKKGYDVKTDMVYLQTDFNLKRKTKTELLPOMGMNDLKLVQLI 240

Qy 241 AYTDWKETOQDPKGTYWNTNYDMMPKGAMAEVRYADGVPGWTMLVNKEESKPDNV 300

Db 241 AYTDWKETOQDPKGTYWNTNYDMMPKGAMAEVRYADGVPGWTMLVNKEESKPDNV 300

Qy 301 YTPLVKELAQNVEWHPYTRKDALBFFDTGIVFPL 360

Db 301 YTPLVKELAQNVEWHPYTRKDALBFFDTGIVFPL 360

Qy 361 KGIK 364

Db 361 KGIK 364

ALIGNMENTS

RESULT 2
 S59932
 Glycerocephosphodiester phosphodiesterase (EC 3.1.4.46) - Haemophilus influenzae (strain 3)
 N; Alternate names: immunoglobulin D-binding protein; protein D
 C; Species: Haemophilus influenzae
 A; Variety: strain 3640
 C; Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 08-Oct-1999
 C; Accession: S59932; S4734
 R; Song, X.M.; Forsgren, A.; Janson, H.
 Infect. Immun. 63, 696-699, 1995.
 A; Title: The gene encoding protein D (hpdi) is highly conserved among Haemophilus influenzae.
 A; Reference number: S59931; PMID:95122210; PMID:7822043
 A; Accession: S59932
 A; Molecule type: DNA
 A; Residues: 1-364 <SON>
 A; Cross-references: EMBL:235657; NID:9525215; PIDN:CAA84716.1; PID:9525216
 A; Experimental source: strain 3640
 C; Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen

Query Match 77.4%; Score 1916; DB 2; Length 364;
 Best Local Similarity 99.2%; Pred. No. 3.7e-122; MisMatches 0; Indels 0; Gaps 0;
 Matches 361; Conservative 0;

Qy 1 MDPKTLAISLLAAGVLAGSSHHSMANTQMSKIIIAHRSAGYLPEHTLESKALAPA 60
 Db 1 MKLKTLAISLLAAGVLAGSSHHSMANTQMSKIIIAHRSAGYLPEHTLESKALAPA 60
 Qy 61 QOADYLEQDLAMTKGRLVYTHDHFGLTDVAKKEPHRKDGRTYVVDFTLKIQSLE 120
 Db 61 QOADYLEQDLAMTKGRLVYTHDHFGLTDVAKKEPHRKDGRTYVVDFTLKIQSLE 120
 Qy 61 QOADYLEQDLAMTKGRLVYTHDHFGLTDVAKKEPHRKDGRTYVVDFTLKIQSLE 120
 Db 61 QOADYLEQDLAMTKGRLVYTHDHFGLTDVAKKEPHRKDGRTYVVDFTLKIQSLE 120
 Qy 121 MTENPETDKGKQAQYVNPREFPLWKSFRINTFEDIEFQGLEKSTGKRVGIYPPIKAPW 180
 Db 121 MTENPETDKGKQAQYVNPREFPLWKSFRINTFEDIEFQGLEKSTGKRVGIYPPIKAPW 180
 Qy 181 FHHQNGKDAAEETLKVLCYKCYDRTKDMYLTQTFDFNLKRKTELLPONGMDLKLVQLI 240
 Db 181 FHHQNGKDAAEETLKVLCYKCYDRTKDMYLTQTFDFNLKRKTELLPONGMDLKLVQLI 240
 Qy 301 YTPLVKELAQNVETHPYTVRDKALPEFTDVLNKSGATGVFTDFDTGVEFL 360
 Db 301 YTPLVKELAQNVETHPYTVRDKALPEFTDVLNKSGATGVFTDFDTGVEFL 360
 Qy 361 KGIK 364
 Db 361 KGIK 364

RESULT 4

S59934
 Glycerocephosphodiester phosphodiesterase (EC 3.1.4.46) precursor - Haemophilus influenzae
 N; Alternate names: immunoglobulin D-binding protein; protein D
 C; Species: Haemophilus influenzae
 A; Variety: strain Eagan; strain HK695
 C; Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
 C; Accession: S59934; S59935; S47336; S47337
 R; Song, X.M.; Forsgren, A.; Janson, H.
 Infect. Immun. 63, 696-699, 1995.
 A; Title: The gene encoding protein D (hpdi) is highly conserved among Haemophilus influenzae.
 A; Reference number: S59931; PMID:95122210; PMID:7822043
 A; Molecule type: DNA
 A; Accession: S59934
 A; Cross-references: UNIPROT:Q06282; EMBL:235659; NID:9525219; PIDN:CAA84718.1; PID:9525222

A; Molecule type: DNA
 A; Residues: 1-364 <SON>
 A; Experimental source: strain Eagan

A; Cross-references: EMBL:Z35660; NID:9525221; PIDN:CAA84719.1; PID:9525222
 A; Cross-references: EMBL:Z35660; NID:9525221; PID:9525222
 A; Experimental source: strain HK95
 C; Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen

F;1-18/Domain: signal sequence #status predicted <SIG>
 F;19-364/Product: protein D #status predicted <MAT>

RESULT 3
 S59931
 Glycerocephosphodiester phosphodiesterase (EC 3.1.4.46) - Haemophilus influenzae (strain 3)
 N; Alternate names: IgD-binding protein; protein D
 C; Species: Haemophilus influenzae
 A; Variety: strain 3639
 C; Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 08-Oct-1999
 C; Accession: S59931; S47333
 R; Song, X.M.; Forsgren, A.; Janson, H.
 Infect. Immun. 63, 696-699, 1995.
 A; Reference number: S59931; PMID:95122210; PMID:7822043
 A; Accession: S59931
 A; Status: preliminary
 A; Molecule type: DNA
 A; Residues: 1-364 <SON>
 A; Cross-references: EMBL:235656; NID:9525213; PIDN:CAA84715.1; PID:9525214
 A; Experimental source: strain 3639
 C; Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen

Query Match 77.3%; Score 1914; DB 2; Length 364;
 Best Local Similarity 99.2%; Pred. No. 3.7e-122; MisMatches 0; Indels 0; Gaps 0;
 Matches 361; Conservative 0;

Qy 1 MDPKTLAISLLAAGTAGCSSHSMANTQMSKIIIAHRSAGYLPEHTLESKALAPA 60
 Db 1 MKLKTLAISLLAAGTAGCSSHSMANTQMSKIIIAHRSAGYLPEHTLESKALAPA 60
 Qy 61 QQADYLEQDLAMTKGRLVYTHDHFGLTDVAKKEPHRKDGRTYVVDFTLKIQSLE 120
 Db 61 QQADYLEQDLAMTKGRLVYTHDHFGLTDVAKKEPHRKDGRTYVVDFTLKIQSLE 120
 Qy 121 MTENPETDKGKQAQYVNPREFPLWKSFRINTFEDIEFQGLEKSTGKRVGIYPPIKAPW 180
 Db 121 MTENPETDKGKQAQYVNPREFPLWKSFRINTFEDIEFQGLEKSTGKRVGIYPPIKAPW 180
 Qy 181 FHHQNGKDAAEETLKVLCYKCYDRTKDMYLTQTFDFNLKRKTELLPONGMDLKLVQLI 240
 Db 181 FHHQNGKDAAEETLKVLCYKCYDRTKDMYLTQTFDFNLKRKTELLPONGMDLKLVQLI 240
 Qy 301 YTPLVKELAQNVETHPYTVRDKALPEFTDVLNKSGATGVFTDFDTGVEFL 360
 Db 301 YTPLVKELAQNVETHPYTVRDKALPEFTDVLNKSGATGVFTDFDTGVEFL 360
 Qy 361 KGIK 364
 Db 361 KGIK 364

Db	121	MTENFETKDGKAQAVQVNPRLWKSFRHTPEDEIEPFIQGLBKSTGKVGYIPEIKAPW	180		RESULT 6
Qy	181	FHHQNGKDIAAETLKVKKYGYDKKCTDMVYIQTDFNBLKRKTELLPQMNDLKLVQLI	240	S59933	glycerophosphodiester phosphodiesterase (EC 3.1.4.46) - Haemophilus influenzae (strain N)
N;Alternate names: immunoglobulin D-binding protein; protein D		N;Alternative names: immunoglobulin D-binding protein; protein D			
Db	181	FHHQNGKDIAAETLKVKKYGYDKKCTDMVYIQTDFNBLKRKTELLPQMNDLKLVQLI	240	C;Species: Haemophilus influenzae	
Qy	241	AYTDWKETQEKDPKGTVNNYNDWMFKPGAMAEVVKYADGVGPWTMLVNKEESKPDNV	300	A;Variety: strain 6-7626	
C;Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 08-Oct-1999		C;Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 08-Oct-1999			
Db	241	AYTDWKETQEKDPKGTVNNYNDWMFKPGAMAEVVKYADGVGPWTMLVNKEESKPDNV	300	C;Accession: S59933; S47335	
R;Song, X.M.; Forsgren, A.; Janson, H.		R;Song, X.M.; Forsgren, A.; Janson, H.			
Infect. Immun. 63, 696-699, 1995		Infect. Immun. 63, 696-699, 1995			
Qy	301	YTPLVKELAQINVEVHPYTVRDLPEFTDYNQMDALLNKSGATGVFDPTCGVFL	360	A;Title: The gene encoding protein D (hpd) is highly conserved among Haemophilus influen	
A;Reference number: S59931; MUID:9512210; PMID:7822043		A;Reference number: S59931; MUID:9512210; PMID:7822043			
Db	301	YTPLVKELAQINVEVHPYTVRDLPEFTDYNQMDALLNKSGATGVFDPTCGVFL	360	A;Accession: S59933	
Qy	361	KGIK 364		A;Molecule type: DNA	
Db	361	KGIK 364		A;Residues: 1-364 <SON>	
C;Cross-references: EMBL:Z35658; NID:9525217; PID:CAA8471.1; PID:9525218		A;Experimental source: strain 6-7626			
C;Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen		C;Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen			
Query Match	76.1%	Score 1885; DB 2; Length 364;			
Best Local Similarity	97.5%	Pred. No. 3-4e-120; Mismatches 4; Indels 5; Gaps 0;			
Matches 355; Conservative					
Qy	1	MDPKTIALSLLAAGTLAGCSSHSMANTOMKSDFKIIIAHRSAGSYLPEHTLESKALAF	60		
Db	1	MRLKLTIALSLLAAGTLAGCSSHSMANTOMKSDFKIIIAHRSAGSYLPEHTLESKALAF	60		
C;Accession: G61086					
C;Species: Haemophilus influenzae					
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004					
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkegaard, P.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.; Science 269, 496-512, 1995					
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A;Title: Whole genome random sequencing and assembly of Haemophilus influenzae Rd.					
A;Reference number: A64000; MUID:95350630; PMID:7542800					
A;Status: nucleic acid sequence not shown; translation not shown					
Db	121	MENFFTKDQKAQTYPNRFLPKWHSFRHTPEDEBPIQLEKSTGKVGYIPEIKAPW	180		
A;Accession: G61086					
A;Molecule type: DNA					
A;Cross-references: UNIPROT:Q06282; GB:U32751; GB:U42023; NID:93212200; PID: AAC22348.1;					
C;Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen					
Query Match	76.8%	Score 1902; DB 2; Length 364;			
Best Local Similarity	98.4%	Pred. No. 2-4e-121; Mismatches 2; Indels 0; Gaps 0;			
Matches 358; Conservative					
Qy	1	MDPKTIALSLLAAGTLAGCSSHSMANTOMKSDFKIIIAHRSAGSYLPEHTLESKALAF	60		
Db	1	MKLKTLIALSLLAAGTLAGCSSHSMANTOMKSDFKIIIAHRSAGSYLPEHTLESKALAF	60		
C;Accession: G61086					
C;Species: Haemophilus influenzae					
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004					
R;Song, X.M.; Forsgren, A.; Janson, H.					
Infect. Immun. 63, 696-699, 1995					
Db	61	QADYLDQDLMTKGRLVYTHDHLFGLTDVAKCPHRHRKDGYYVIDFTLKETOSL	120		
Db	61	QADYLDQDLMTKGRLVYTHDHLFGLTDVAKCPHRHRKDGYYVIDFTLKETOSL	120		
C;Accession: G61086					
C;Species: Haemophilus influenzae					
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004					
R;Song, X.M.; Forsgren, A.; Janson, H.					
Infect. Immun. 63, 696-699, 1995					
Db	121	YTPLVKELAQINVEVHPYTVRDLPEFTDYNQMDALLNKSGATGVFDPTCGVFL	360		
Db	301	YTPLVKELAQINVEVHPYTVRDLPEFTDYNQMDALLNKSGATGVFDPTCGVFL	360		
C;Accession: G61086					
C;Species: Haemophilus influenzae					
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004					
R;Song, X.M.; Forsgren, A.; Janson, H.					
Infect. Immun. 63, 696-699, 1995					
Db	121	MTENFETKDGKAQAVQVNPRLWKSFRHTPEDEIEPFIQGLEKSTGKVGYIPEIKAPW	180		RESULT 7
Db	121	MTENFETKDGKAQAVQVNPRLWKSFRHTPEDEIEPFIQGLEKSTGKVGYIPEIKAPW	180	S59936	glycerophosphodiester phosphodiesterase (EC 3.1.4.46) - Haemophilus influenzae (strain N)
N;Alternate names: IgD-binding protein; protein D		N;Alternative names: IgD-binding protein; protein D			
Db	181	FHHQNGKDIAAETLKVKKYGYDKKCTDMVYIQTDFNBLKRKTELLPQMNDLKLVQLI	240	C;Species: Haemophilus influenzae	
Db	181	FHHQNGKDIATETLKVKKYGYDKKCTDMVYIQTDFNBLKRKTELLPQMNDLKLVQLI	240	A;Variety: strain NTC 8468	
C;Accession: S59936; S47338		C;Accession: S59936; S47338			
R;Song, X.M.; Forsgren, A.; Janson, H.		R;Song, X.M.; Forsgren, A.; Janson, H.			
Infect. Immun. 63, 696-699, 1995		Infect. Immun. 63, 696-699, 1995			
Db	241	AYTDWKETQEKDPKGTVNNYNDWMFKPGAMAEVVKYADGVGPWTMLVNKEESKPDNV	300	A;Title: The gene encoding protein D (hpd) is highly conserved among Haemophilus influen	
Db	241	AYTDWKETQEKDPKGTVNNYNDWMFKPGAMAEVVKYADGVGPWTMLVNKEESKPDNV	300	A;Reference number: S59931; MUID:9512210; PMID:7822043	
C;Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen		C;Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen			
Qy	301	YTPLVKELAQINVEVHPYTVRDLPEFTDYNQMDALLNKSGATGVFDPTCGVFL	360	A;Molecule type: DNA	
Db	301	YTPLVKELAQINVEVHPYTVRDLPEFTDYNQMDALLNKSGATGVFDPTCGVFL	360	A;Residues: 1-364 <SON>	
Db	361	KGIK 364		A;Cross-references: EMBL:Z35661; NID:9525222; PID:CAA84720.1; PID:9525224	
Db	361	KGIK 364		A;Experimental source: strain NTC 8468	
Db	361	KGIK 364		C;Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen	

Query Match	76.1%;	Score 1885;	DB 2;	Length 364;		Qy	244 DWKETQEKDPKGYWNNYDWMFKPGAMAEVVVKYADGVPGWYMLVNKEESKPDMIVYTP 303
Best Local Similarity	97.3%;	Pred. No.	3.4e-120;			Db	240 DNNEEQKQDPGSWNNYDWMFKPGAMQVAEYDGTGDPYHMLI-EETSQPNIKLRTG 298
Matches	355; Conservative	2;	Mismatches	0;	Gaps	0;	
Qy	1 MDPKTLAISLLAAGVLAGCSSHSSNMANOMSKDKIIIAHRGASGLPHEHTLESKALATA 60		Qy	304 LVKELAQYNNVBHPVTVRKDPEFTDYNQMDALLNSGATGVFTDPEPDGVEFL 360			
Db	1 MKLTKTLAISLLAATGGVCSHHSSNMANOMSKDKIIIAHRGASGLPHEHTLESKALATA 60		Db	299 MVQDQAQNKLUVHPVTVRSDKLPETTDVNQLDLYNKAGVNGLFTDFDKAVKPL 355			
Qy	61 QADYLQDQDLANTKDGRLVUVTHDHELDGLTDVAKKEPHRKDGRRYYVDEFTLKEIOSLE 120						
Db	61 QADYLQDQDLANTKDGRLVUVTHDHELDGLTDVAKKEPHRKDGRRYYVDEFTLKEIOSLE 120						
Qy	121 MTENFETKDGKQAAQYPNRPLWKSFRIRTFEDELFICQLEKSTGKVGKYPPIKAPW 180		Qy	121 MTENFETKDGKQAAQYPNRPLWKSFRIRTFEDELFICQLEKSTGKVGKYPPIKAPW 180			
Db	121 MTENFETKDGKQAAQYPNRPLWKSFRIRTFEDELFICQLEKSTGKVGKYPPIKAPW 180						
Qy	181 FHHQNGKDIAAETKLVKLYKGDKCTDMVYLQTDFDENELKRKTEELLPGMGLKLXVQLI 240		Qy	181 FHHQNGKDIAAETKLVKLYKGDKCTDMVYLQTDFDENELKRKTEELLPGMGLKLXVQLI 240			
Db	181 FHHQNGKDIAAETKLVKLYKGDKCTDMVYLQTDFDENELKRKTEELLPGMGLKLXVQLI 240						
Qy	241 AYTDWKETQEKDPKGYWNNYDWMFKPGAMAEVVVKYADGVPGWYMLVNKEESKPDMIV 300		Qy	241 AYTDWKETQEKDPKGYWNNYDWMFKPGAMAEVVVKYADGVPGWYMLVNKEESKPDMIV 300			
Db	241 AYTDWKETQEKDPKGYWNNYDWMFKPGAMAEVVVKYADGVPGWYMLVNKEESKPDMIV 300						
Qy	301 YTPLUKBLAQINNEVHPYTVRKDALPEFTDYNQMDALLNSGATGVFTDPEPDGVEFL 360		Qy	301 YTPLUKBLAQINNEVHPYTVRKDALPEFTDYNQMDALLNSGATGVFTDPEPDGVEFL 360			
Db	301 YTPLUKBLAQINNEVHPYTVRKDALPEFTDYNQMDALLNSGATGVFTDPEPDGVEFL 360						
Qy	361 KGIK 364		Qy	361 KGIK 364			
Db	361 KGIK 364						
Db	*						
RESULT 8							
D91019	9 glycerophosphodiester phosphodiesterase [imported] - Escherichia coli (strain O157:H7, S		Qy	4 KTLAISLLAAGVLAGCSSHSSNMANOMSKDKIIIAHRGASGLPHEHTLESKALAFQAQA 63			
C;Species: Escherichia coli			Db	6 KNLSNAIMMSLVMGSSAAMD----SNEKIVTAHRGASGLPHEHTLPKAMAYAQQA 59			
C;Accession: D91019	C;Sequence revision 18-Jul-2001 #text_change 09-Jul-2004		Qy	64 DYLEQDLAMTKDGRLVVHDHFLDGLTDVAKKPFHHRKOGRRYYVDFLTKEIOSLEMTE 123			
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Han, C.G.			Db	60 DYLEQDLVMTKDDHVVLDHYLDRVTDAFRFPRAKDGRYAIDFTLEISLKFT 119			
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.			Qy	124 NFETDKGKOAQYVNPREFLWKSHPRIHTEDEFIPTGLEKSTGKVGKYPEIKAPWFHH 183			
DNA Res. 8.11-22 2001	A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene		Db	120 GFDIENGKRVQTYPERFMGKSDPRVHTFEEIEVQGLNHSIGNIGNIYPEIKAPWFHH 179			
A;Reference number: A99629; PMID:21156231; PMID:11258796	A;Accession: D91019		Qy	184 QNGKUAAETKLVKLYKGDKCTDMVYLQTDFDENELKRKTEELLPGMGLKLVOLIAYT 243			
A;Status: preliminary	A;Molecule type: DNA		Db	180 QEGKIAAKTLEVYKYGKGDVKYLOCDFADSLKRKNELERGMHDNLVOLIAYT 239			
A;Residues: 1-358 <RAY>	A;Cross-references: UNIPROT:Q8XE17; GB:BA000007; PID:BAB36547_1; PID:gi13362594; GSPDB:G		Qy	244 DWKETQEKDPKGYWNNYDWMFKPGAMAEVVVKYADGVPGWYMLVNKEESKPDMIVYTP 303			
A;Experimental source: strain O157:H7, substrate RIMD 0509952	A;Genetics:		Db	240 DNNEEQKQDPGSWNNYDWMFKPGAMQVAEYDGTGDPYHMLI-EETSQPNIKLRTG 298			
A;Gene: EC3124			Qy	304 LVKELAQYNNVBHPVTVRKDPEFTDYNQMDALLNSGATGVFTDPEPDGVEFL 360			
Query Match	50.9%;	Score 1261.5;	DB 2;	Length 358;		Db	299 MVQDQAQNKLUVHPVTVRSDKLPETTDVNQLDLYNKAGVNGLFTDFDKAVKPL 355
Best Local Similarity	64.7%;	Pred. No.	5.3e-78;				
Matches	231; Conservative	55;	Mismatches	64;	Gaps	2;	
Qy	4 KTLAISLLAAGVLAGCSSHSSNMANOMSKDKIIIAHRGASGLPHEHTLESKALAFQAQA 63						RESULT 10
Db	6 KNLSNAIMMSLVMGSSAAMD----SNEKIVTAHRGASGLPHEHTLPKAMAYAQQA 59						S15945
Qy	64 DYLEQDLAMTKDGRLVVHDHFLDGLTDVAKKPFHHRKOGRRYYVDFLTKEIOSLEMTE 123						glycerophosphodiester phosphodiesterase (EC 3.1.4.46) precursor, periplasmic - Escherich
Db	60 DYLEQDLVMTKDDHVVLDHYLDRVTDAFRFPRAKDGRYAIDFTLEISLKFT 119						C;Species: Escherichia coli
Qy	124 NFETDKGKOAQYVNPREFLWKSHPRIHTEDEFIPTGLEKSTGKVGKYPEIKAPWFHH 183						C;Accession: S15945; S00871; E64984; S72653; S14522
Db	120 GFDIENGKRVQTYPERFMGKSDPRVHTFEEIEVQGLNHSIGNIGNIYPEIKAPWFHH 179						R;Tommassen, J.; Englmeier, K.; Cole, S.T.; Overduin, P.; Larson, T.J.; Boos, W.
Qy	184 QNGKDTAAETKLVKLYKGDKCTDMVYLQTDFDENELKRKTEELLPGMGLKLVOLIAYT 243						Mol. Gen. Genet. 246, 321-327, 1991
Db	180 QEGKIAAKTLEVYKYGKGDVKYLOCDFADSLKRKNELERGMHDNLVOLIAYT 239						A;Title: Characterization of two genes, gfpQ and gfpO, encoding glycerophosphoryl dieste
Qy							A;Reference number: S15945; PMID:91238712; PMID:1851953
Db							A;Molecule type: DNA
A;Residues: 1-358 <TOM>	A;Cross-references: UNIPROT:D09394; EMBL:X56907; NID:941580; PID:CAA40223_1; PID:g41581						R;Englmeier, K.; Boos, W.; Cole, S.T.

Mol: Microbiol. 1, 251-258, 1987 A; Title: Nucleotide sequence and transcriptional startpoint of the <i>glpT</i> gene of Escherichia coli K-12. A; Reference number: Sc00868; MUID:88201663; PMID:3329281	A;Title: Complete genome sequence of a multiple drug resistant <i>Salmonella enterica</i> serovar A;Reference number: AB0502; MUID:21534347; PMID:11677608 A;Accession: AH0791 A;Status: preliminary A;Molecule type: DNA A;Residues: 1-9 <BIG> A;Cross-references: EMBL:Y00536; NID:941586; PIDN:CAA68599_1; PID:9581099 R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co .A.; Rose, D.J.; Mau, B.; Shao, Y. Science 277, 1453-1452, 1997 A;Title: The complete genome sequence of <i>Escherichia coli</i> K-12. A;Reference number: A64720; MUID:97426617; PMID:9278503 A;Accession: E61994 A;Status: preliminary; nucleic acid sequence not shown; translation not shown A;Molecule type: DNA A;Residues: 1-358 <BLAT> A;Cross-references: GB:AE000314; GB:U00096; NID:91788570; PIDN: AAC75299_1; PID:91788572; A;Experimental source: strain K-12, substrate MG1655 R;Gonzalez-Gil, G.; Brinckmann, P.; Kahmann, R. Mol: Microbiol. 22, 21-29, 1996 A;Title: Fts is a regulator of metabolism in <i>Escherichia coli</i> . A;Reference number: S72651; MUID:97055418; PMID:8859705 A;Accession: S72653 A;Molecule type: Protein A;Residues: 26, 'I', 28-44 <GON> C;Genetics:	Query Match 49 8%; Score 1233; DB 2; Length 356; Best Local Similarity 64.5%; Pred. No. 4.5e-76; Matches 231; Conservative 47; Mismatches 70; Indels 10; Gaps 2; Qy 4 KTTLASLLAAGVLAGCSSHSSNNMANTQMSKSPRIITIAHRCGGLPHTLESKALAFAAQA 63 Db 6 KNLSVALMLAGMTIGSGA-----VAAEKVDTIAHRCGGLPHTLPKAMAYAQA 56
	Qy 64 DYLEDQDLAMTKDGRVVYIHDHFLDGLTDIVAKKKFPHHRKGDRYYVDFLTKEIOSLEMT 123 Db 57 DYLEDQDLAMTKDGRVVYIHDHFLDGLTDIVAKKKFPHHRKGDRYYVDFLTKEIOSLEMT 116 Qy 124 NFETRDGKQAOVYPNRPRPLKSHPRHTFEDIEFTQGLEKSTGKVGIVPEIKAPWFAH 183 Db 117 GFIDIENGKKVQTQYTPSRFPMGSKSDFRHTFEEETFVQGLNHSTGNIGVPEIKAPWFAH 176 Qy 184 QNGKDAAEETLKVKKYGYDKRTDVKVLYQDFTDNEKLRKTIKELLPGMGMIDKLVQJLJAYT 243 Db 177 QEGKDAAKTLEVLKVKYGYGKQDQTVYLQCPDVAEKLKRKINELPQMGMIDNLVQJLJAYT 236 Qy 244 DWKETQKDPKGKVMWVNYNDWMFKGAMAEVVKAQYADGVGPCKWMLVNKEESKPDNTVYTP 303 Db 237 DNNETQQKOPGPGRWVNYNDWMFKGAMAEVVKAQYADGVGPCKWMLVNKEESKPDNTVYTP 303 Qy 304 LYKELAQNVNVBHYPTVRDALPEFTDYNQMDALLNSCATGVTDFDPDTGVEFLK 361 Db 296 MVQDAHQNKMVNVBHYPTVRQDLPDFTDVNQLYDLYNKAQVDGLFTDFPKAVNFLQ 353	
	RESULT 12 AC0466 glycerophosphodiester phosphodiesterase (EC 3.1.4.46) (imported) - <i>Yersinia pestis</i> (stra C;Species: <i>Yersinia pestis</i> C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004 R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dugan, G.; Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001 A;Title: Genome sequence of <i>Yersinia pestis</i> , the causative agent of plague. A;Reference number: AB0001; MUID:21470413; PMID:1586360 A;Accession: AC0466 A;Status: preliminary A;Molecule type: DNA A;Residues: 1-371 <KUR> A;Cross-references: UNIPROT:Q8ZAH4; GB:AL590842; A;Gene: gipQ C;Keywords: phosphoric diester hydrolase	Query Match 48 8%; Score 1210; DB 2; Length 371; Best Local Similarity 62.7%; Pred. No. 1.7e-74; Matches 227; Conservative 54; Mismatches 75; Indels 6; Gaps 4;
	Qy 64 DYLEDQDLAMTKDGRVVYIHDHFLDGLTDIVAKKKFPHHRKGDRYYVDFLTKEIOSLEMT 123 Db 60 DYLEDQDLAMTKDNLVYIHDHFLDGLTDIVAKKKFPHHRKGDRYYVDFLTKEIOSLEMT 119 Qy 124 NFETRDGKQAOVYPNRPRPLKSHPRHTFEDIEFTQGLEKSTGKVGIVPEIKAPWFAH 183 Db 120 GDPDIENKKVQTQYTPGRPMGSKSDFRHTFEEETFVQGLNHSTGNIGVPEIKAPWFAH 179 Qy 184 QNGKDAAEETLKVKKYGYDKRTDMYLQDFNELLRKTIKELLPGMGMIDKLVQJLJAYT 243 Db 180 QEGKDAAKTLEVLKVKYGYGKQDQTVYLQCPDVAEKLKRKINELPQMGMIDNLVQJLJAYT 239 Qy 244 DWKETQKDPKGKVMWVNYNDWMFKGAMAEVVKAQYADGVGPCKWMLVNKEESKPDNTVYTP 303 Db 240 DNNETQQKOPGPGRWVNYNDWMFKGAMAEVVKAQYADGVGPCKWMLVNKEESKPDNTVYTP 303 Qy 304 LYKELAQNVNVBHYPTVRDALPEFTDYNQMDALLNSCATGVTDFDPDTGVEFLK 360 Db 299 MVQDAHQNKMVNVBHYPTVRQDLPDFTDVNQLYDLYNKAQVDGLFTDFPKAVNFLQ 355	Qy 4 KTTLASLLAAGVLAGCSSHSSNNMANTQMS--DKIIIAHRCGASYLPEHTLESKALAFAA 60 Db 6 KTTLASLLAAGVLAGCSSHSSNNMANTQMS--DKIIIAHRCGASYLPEHTLESKALAFAA 64 Qy 61 QGADTLEQDIAKTMKGRLVYIHDHFLDGLTDPKAVFLKREIQSLE 120 Db 65 QGADTLEQDIAKTMKGRLVYIHDHFLDGLTDPKAVFLKREIQSLE 124 A;Authors: Party, C.; Quail, M.; Skelton, J.; Stevens, K.
	RESULT 11 AH0791 9-glycerophosphodiester phosphodiesterase (EC 3.1.4.46) - <i>Salmonella enterica</i> subsp. enter C;Species: <i>Salmonella enterica</i> subsp. enterica serovar Typhi A;Note: this species has also been called <i>Salmonella typhi</i> C;Accession: A610791 R;Parkhill, J.; Dougan, G.; James, K.D.; Pickard, D.; Wain, J.; Churcher, .S.; Connor, P.; Cronin, A.; Davis, P.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P. Nature 413, 848-852, 2001 A;Authors: Party, C.; Quail, M.; Skelton, J.; Stevens, K.	Query Match 48 8%; Score 1210; DB 2; Length 371;

RESULT 13									
F71346									
Qy	121	MTEFNE-TKDGRDQAQVYNRFLPKWSHFRHTFEDEIFIQLEKSTGKVGITYPEIKAP	179						
Ddb	125	FTECFDIDKGKKVQSPNRFPMGKSDFRVTFOQEETFIQLNASTGRTGITYPEIKAP	184						
Qy	180	WFHHQNGKDIAAETKLKVKKYGYDKKTDMLYLQTDFDNELKRKTBLLPOMGMDLKLVLQ	239						
Ddb	185	WFHHQEGDIDSTKLVEKEYGTNKDKVYIQLSFVNELLRKINPLERPKGMGDLKLVLQ	244						
Qy	240	IAYTDWKETQEDPKGTVWNWNYNDWMFKPGAMAEVVKYADGVGPWYMLVNKEESKPDN	299						
Ddb	245	IAYTDWNETYEQLPDGKWMWNYNDWMFKPGAMKEVARYADGJPDPYHMLVYKTSFAKD-I	303						
Qy	300	VYTTELVKELAQTVNEVHYTVDKDALBEFTDVNQMYDALLINKSGATGTVFDPPDGVEF	359						
Ddb	304	KLTDLVKEAHASNMVNHPTIRADRPLPKYATDINQLEFDITYNQAQDGVFTDFPQGIQF	363						
Qy	360	LK 361							
Ddb	364	LQ 365							
RESULT 14									
Qy	probable glycerophosphorylilester phosphodiesterase (g1bQ) - syphilis spirochete								
C;Species:	Treponema pallidum subsp. Pallidum (syphilis spirochete)								
C;Date:	24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004								
C;Accession:	F71346								
C;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin, J.K.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDaniel, L.; Weidman, H.O.; Venter, J.C.									
Science, 281, 375-388, 1998									
A;Title:Complete genome sequence of Treponema pallidum, the syphilis spirochete.									
A;Accession: F71346									
A;Status: preliminary; nucleic acid sequence not shown; translation not shown									
A;Molecule type: DNA									
A;Accessories: 1-356 <C02>									
A;Cross-references: UNIPROT:030405; GB:AE001207; PMID:9832770; MUID:9665876									
A;Experimental source: strain Nichols									
C;Genetics:									
A;Gene: TP0257									
Query	Match	43-44;	Score	1074;	DB	2;	Length	356;	
Best	Local	Similarity	54.64;	Pred.	No.	2.56-65;			
Matches	190;	Conservative	67;	Mismatches	79;	Indels	12;	Gaps	1;
Qy	13	AGVIAAGCSSSHSSNMANTOMKSDDKIIIAHRGAGSYLPLPHTLSKALLFAQQADYLSDQDLAM	72						
Ddb	15	AAVAGCA-----SRRMTVAYRGAGYVPEHTFLSKVLAFAQGADYLQODVV	62						
Qy	73	TKLTKYGYDKKIDMLYLQTDPDNELKRKTBLLPOMGMDLKLVOLIAYTDWKETQEKD	132						
Ddb	63	SKRNQLIQAQSHTLDNNMVAEKFPIRQADGFHYVIDFTIVEELSLLRATNSFYTRGRKH	122						
Qy	133	AQYVNPNEFLPKWSHFRHTFEDEIFIQLEKSTGKVGITYPEIKAPWFHFFHQNGKDI	192						
Ddb	123	TPYQGQRPLWPKCQFRHTFEERLQFTRGLETTGKIGIYSEIKVWPWFFHQEGKDI	182						
Qy	193	TLALKYGYQSRSQSLVYQTDIFNELLKRKRELLPKYENAVKLQRVAYTDQRETKQEKD	252						
Ddb	183	PKGYWVNWNYNDWMFKPGAMAEVVKYADGVGPWYMLVNKEESKPDNIVTPVLYKELAQXN	312						
Qy	253	SRGWVINTNWNMNEFPGNQKTAKYADGVPMRMLIENENQSVGAVRLSMVSIAQDAK	302						
Ddb	243	VEHPYTYTVDKALPEFFTDDVNQMYDALLINKSGATGVFTDFPDTGVBF	360						
Qy	303	LECHVHTVRKETLPSYARTMDMENSLUFKQTSANVVLTDFFPLGKFL	350						

Query Match Score 425; DB 2; Length 1027;
 Best Local Similarity 29.4%; Pred. No. 9.6e-21;
 Matches 137; Conservative 66; Mismatches 147; Indels 116; Gaps 20;

Qy 371 IIAHGRASGYLPBPHTLSKALLAFAQQADYLBODLAMTKDGRALVVIHD-----H 84
 Db 180 VIGHRGASGFREHTLESYKLAIEQGADPFLATYKDGVLJARHEPALAVLNADGSVN 239

Qy 85 FUDGLTDYAK--KFPHHRIRK---DGR---YYVIDFTLKEIOSLENTENFETKDGRQAAQV 135
 Db 240 FSITTTTNYQIAKPSDRKTVNLDGTETGWAEDFTLAEIKELRAIERL----- 289

Qy 136 YPNRFPLWKSFRHFRHTPEDEIEPIQGLEKSTGKVGYPEIKAQWPFTHQ----- 184
 Db 290 -PFRDOSTNGQFTIPTLAELIDLVKQVBAETSKKKIGYPETKHPYTAQEATYVGTTTEKI 348

Qy 185 --NGKDIAAETLKVLUKKGYDKKTD--MVLQLQTFDFNELLKRKTELLPQMGMDLKUQVL 240
 Db 349 NRNISQILIDTLKA----NNFTDPSRIFQSPEVNLKELHTIMPAGVDIPLVQLF 402

Qy 241 AYTDWKEEQERDKDPKGYWWNN-YDMMEK-----PGAMAEVVKYADGVGPWYML 288
 Db 403 -----DAVIDDINGRLLETRPYDFIVSDFRTYGDLRTPAGLAETAEYAGIGGPWKRM 456

Qy 289 VNKEESKPDN-----IVYTPLYKELAQQNVBVHPYTVRKDALPFP 328
 Db 457 VSVRGTDANNQGADDYNGDGAVNDADKTLIPPTTLVQDAHNVGLOVHPYTFRDE--ER 513

Qy 329 FTDVNON-----YDALLINKSATGVTDFPDTGVEFLGKIKSMDGGRAGVALVR-DYKL 382
 Db 514 YLAANTYQGNPELEYQQLF-QLGVDALFTDFPIT-ADRVRDRSLSPNN--IVRSP-- 564

Qy 383 YNKNSSSNSTLURNLGERHARAMDGGKRAVALVRSDFKFYEDANGT 428
 Db 565 QNPDVLSGDAFANLG---GSRGEFGG---AINASKFKLYMLLEG 603

Search completed: June 22, 2005, 11:41:29
 Job time : 43 secs

THIS PAGE BLANK (USPTO)

Scoring table:	BLOSUM62	ALIGMENTS			
Gapop:	10.0 , Gapext: 0.5				
Searched:	1612378 seqs, 512079187 residues				
Total number of hits satisfying chosen parameters:	1612378				
Minimum DB seq length:	0				
Maximum DB seq length:	2000000000				
Post-processing:	Minimum Match 0% Maximum Match 100% Listing first 45 summaries				
Database :	UniProt_03-* 1: uniprot_sprot: 2: uniprot_trembl: *				
Prey.	No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.				
SUMMARIES					
Result No.	Score	Query Match	Length	DB ID	Description
1	1902	76.8	364	1 GLPQ_HAEIN	09c102 haemophilus
2	1577	63.7	358	2 Q8LQ04	09c104 pasteurilla
3	1501.5	60.6	368	2 Q65R11	Q65R11 manneheimia
4	1498	60.5	357	2 Q8RHT6	Q8RHT6 fusobacteri
5	13.63	55.0	309	2 QTP6W6	Q8xe17 escherichia
6	1261.5	50.9	358	2 Q8XB17	Q8cvv6 escherichia
7	1251.5	50.6	358	2 GLPQ_ECOLI	P09394 escherichia
8	1251.5	50.5	358	1 Q7UC72	Q7uc72 shigella fl
9	1551.5	50.5	358	2 Q83QU2	Q83QU2 shigella fl
10	1218	50.4	377	2 Q6cz14	Q6cz14 erwinia car
11	124.8	50.4	361	2 Q82NG7	Q8ang7 salmonella
12	1235	49.9	356	2 Q82556	Q82556 salmonella
13	1233	49.8	356	2 Q7mzy7	Q7mzy7 photorhabdu
14	1210.5	48.9	358	2 Q76FX7	Q66fx7 yersinia pe
15	1210	48.8	371	2 Q8ZAH4	Q8zah4 yersinia pe
16	1210	48.8	371	2 Q8D1J5	Q8d1j5 yersinia pe
17	1210	48.8	372	2 Q94X5U6	Q9x5u6 treponema p
18	1082	43.7	356	2 GLPQ_TREPBA	Q30405 treponema p
19	1074	43.4	356	1 Q71U57	Q71u57 treponema s
20	1074	43.4	356	2 Q71U58	Q71u58 treponema p
21	1074	43.4	356	2 Q71U59	Q71u59 treponema p
22	1074	43.4	356	2 Q71U64	Q71u64 treponema p
23	1074	43.4	356	2 Q7QMY5	Q7qmy5 anophelles g
24	1072.5	43.3	341	2 Q61UT65	Q61ut65 photobacter
25	1026	41.4	360	2 Q87M77	Q87m77 vibrio para
26	987.5	39.5	351	2 Q7MDR9	Q7mdr9 vibrio vuln
27	979	39.5	352	2 QBD6W7	Qbd6w7 vibrio vuln
28	979	39.5	352	2 Q9KN30	Q9kn30 vibrio chol
29	939.5	38.6	371	2 Q9F8D4	Q9f8d4 borrelia tu
30	839.5	33.9	338	2 Q9F8D3	Q9f8d3 borrelia co
31	836	33.8	337	2 Q8YY04	Q8yy04 anabaena sp

RESULT 1 GLPQ_HAEIN		
ID	GLPQ_HAEIN	STANDARD;
AC	0061202;	PRT; 364 AA.
DT	01-NOV-1995 (Rel. 32, Created)	
DT	01-NOV-1995 (Rel. 32, Last sequence update)	
DT	25-OCT-2004 (Rel. 45, Last annotation update)	
DB	Glycerophosphoryl diester phosphodiesterase precursor (EC 3.1.4.46)	
DB	(Glycerophosphodiester phosphodiesterase) (Surface-exposed lipoprotein protein).	
DB	(Protein D) (Immunoglobulin D-binding protein) (IgD-binding protein).	
GN	Name=glpQ; Synonyms=hpd; OrderedLocusNames=II0689;	
OS	Haemophilus influenzae.	
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;	
OX	NCBI_TaxID=727;	
RN	[1]	
RP	SEQUENCE FROM N.A. / ATCC 51907; SEQUENCE FROM N.A. / NCBI_TaxID=727;	
RC	STRAIN=Rd / K920 / ATCC 51907; MEDLINE=953350630; PubMed=7542800;	
RX	RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.B.F., Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M., McKenney K., Sutton G.G., FitzHugh W., Fields C.A., Gocayne J.D., Scott J.F., Shirley R., Liu L.-I., Glodck A., Kelley J.M., Weidman J.F., Phillips C.A., Spriggs T., Heblom E., Cotton M.D., Utterback T.R., Hanna M.C., Nguyen D.T., Sandek D.M., Brandon R.C., Fine L.D., Fritchman J.L., Fuhrmann J.L., Geeseygen N.S.M., Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O., Venter J.C.; "Whole-genome random sequencing and assembly of Haemophilus influenzae Rd"; Sciencie 269:496-512(1995). RN [2]	
RP	SEQUENCE FROM N.A. / ATCC 51907;	
RC	STRAIN=NTH 772;	
RX	RA Jansson H., Heden L.-O., Grubb A., Ruan M., Forsgren A.; "Protein D, an immunoglobulin D-binding protein of Haemophilus influenzae: cloning, nucleotide sequence, and expression in Escherichia coli.", Infect. Immun. 59:119-125(1991). RN [3]	
RP	SEQUENCE FROM N.A. / Serotype B; STRAIN=Mima / Serotype B; MEDLINE=94011360; PubMed=1987023;	
RC	RA Jansson H., Ruan M., Forsgren A.; "Limited diversity of the protein D gene (hpd) among encapsulated and nonencapsulated Haemophilus influenzae strains.", Infect. Immun. 61:4546-4552(1993). RN [4]	
RP	SEQUENCE FROM N.A. / Serotype B; STRAIN=NTH 772;	
RC	RA Jansson H., Ruan M., Forsgren A.; "Protein D gene (hpd) among encapsulated and nonencapsulated Haemophilus influenzae strains.", Infect. Immun. 61:4546-4552(1993). RN [5]	
RP	SEQUENCE FROM N.A. / Serotype B; STRAIN=3639 / 3640, 6-7626, Bagan / Serotype B, HK695 / Serotype B, and NCTC 8468 / Serotype B; NCTC 8468 / Serotype B; MEDLINE=9512210; PubMed=7822043;	
RC	RA Jansson H., Ruan M., Forsgren A.; "Protein D gene (hpd) among encapsulated and nonencapsulated Haemophilus influenzae strains.", Infect. Immun. 61:4546-4552(1993). RN [6]	

RA	Song X.-M., Forssgren A., Jansson H.,	VARIANT	191	191	T -> A (in strain Bagan, strain 3639, strain 6-
RT	"The gene encoding protein D (hpD) is highly conserved among				strain 3640, strain NCTC 8468, strain 6-
RT	Haemophilus influenzae type b and nontypeable strains.;"				strain HK695 and strain Minna).
RL	Infect. Immun. 63:696-699(1995).				P -> S (in strain 6-7626).
[5]					Q -> K (in strain 6-7626).
RP	CHARACTERIZATION.				B -> A (in strain Bagan, strain 3639,
RC	STRAIN=NTHI 772;				strain NCTC 8468, strain 6-7626, strain
RX	MEDLINE=92192801; PubMed=1548059;				HK695 and strain Minna).
RA	Jansson H., Heden L.-O., Forssgren A.;	VARIANT	253	253	A -> V (in strain Bagan, strain 3640,
RT	"Protein D, the immunoglobulin D-binding protein of Haemophilus				strain HK695 and strain Minna).
inf	influenzae, is a lipoprotein.";				K -> E (in strain 6-7626).
RL	Infect. Immun. 60:1336-1342(1992).	VARIANT	310	310	
CC	-1- FUNCION: Glycerophosphoryl diester phosphodiesterase hydrolyzes				
CC	deacylated phospholipids to G3P and the corresponding alcohols.	SEQUENCE	327	327	
CC	Has a specific affinity for human immunoglobulin D myeloma				
CC	protein.				
CC	-1- CATALYTIC ACTIVITY: A glycerophosphodiester + H(2)O = an alcohol +				
CC	sn-glycerol 3-phosphate.				
CC	-1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid				
CC	-1- anchor.				
CC	-1- PTM: Contains both ester- and amide-linked fatty acids.				
CC	-1- MISCELLANEOUS: The sequence shown is that of strains NTHI 772 and				
CC	/ KW20.				
CC	-1- SIMILARITY: Belongs to the glycerophosphoryl diester				
CC	phosphodiesterase Family.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to licensee@isb-sib.ch).				
CC	-----				
DR	EMBL: M32751; AAC22348.1; -;			RESULT 2	
DR	EMBL: M37487; AAA24998.1; -;			Q9CL04	
DR	EMBL: L12445; AAA24999.1; -;			ID Q9CL04	
DR	EMBL: Z33556; CAA84715.1; -;			AC Q9CL04;	
DR	EMBL: Z33567; CAA84716.1; -;			DT 01-JUN-2001 (TREMBLrel. 17, Created)	
DR	EMBL: Z33558; CAA84717.1; -;			DT 01-JUL-2004 (TREMBLrel. 27, Last annotation update)	
DR	EMBL: Z33569; CAA84718.1; -;			DE Giph0.	
DR	EMBL: Z33560; CAA84719.1; -;			GN Name=g120; OrderedLocusName=PM1444;	
DR	EMBL: Z33561; CAA84720.1; -;			OS Pasteurellaceae; Pasteurella.	
PIR	G64086; G64086.			OC Pasteurellales; Pasteurella.	
DR	PIR; S59934; S59934.			OX	
DR	TIGR; H10889; -.			RN	
DR	IntersPro; IPR004129; GDPD.			RP SEQUENCE FROM N.A.	
PFam	PF01009; GDPD_1.			RC STRAIN=Ph70; PubMed=11248100; DOI=10.1073/pnas.051634598;	
DR	PROSITE; PS00013; PROKAR_LIPOPROTEIN_1.			RX RA May B.-J. Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;	
KW	Complete proteome; Glycerol metabolism; Hydrolase; Lipoprotein;			RT RA "Complete genomic sequence of Pasteurella multocida Ph70.";	
Outer membrane; Palmitate; Signal.				RT RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).	
PT	SIGNAL 1 18			RN RN SEQUENCE FROM N.A.	
PT	CHAIN 19 364			RC STRAIN=X-73; PubMed=1473884; DOI=10:1016/j.micinf.2003.09.023;	
PT	LIPID 19 19			RC RA Lo M., Boyce J.D., Wilkie I.W., Adler B.;	
PT	VARIANT 13 13			PT VARIANT 144 144 "Characterization of two lipoproteins in Pasteurella multocida."	
PT	VARIANT 16 16			PT VARIANT 168 168 Microbes Infect. 6:58-67(2004).	
PT	VARIANT 25 25				
PT	VARIANT 28 28				
PT	VARIANT 34 34				
PT	VARIANT 62 62				
PT	VARIANT 63 63				
PT	VARIANT 98 98				
PT	VARIANT 99 99				
PT	VARIANT 144 144				
PT	VARIANT 168 168				

DR	EMBL; AE006181; AAK03528.1; -	Qy	64	DYLEQDLAMTKDQGRLVYIHDHFLDGLTDYAKKPFPHRKDGRTYYTDFTLKEIOSLEMTE 123
DR	EMBL; AY28825; AAOT03332.1; -	Db	65	DYLEQDLAMTKDQGRLVYIHDHFLDGLTDYAKKPFPHRKDGRTYYTDFTLKEIOSLEMTE 124
GO;	GO:0008889; P: glycerophosphodiester phosphodiesterase act. . . ; IEA.	Qy	124	NFETRDKQKQOQVNPRLWKSHPHRHTPEDELEPIQLEKSTGKCGYIYPEIKAPWKH 183
GO;	GO:0006071; P: glycerol metabolism; IEA.	Db	125	NFKTENGKQKQOQVNPRLWKSHPHRHTPEDELEPIQLEKSTGKCGYIYPEIKAPWKH 184
InterPro;	IPR004129; GDDP.	Qy	184	QNGKDIAETLKLKQYGDRTKTDQMDVLYQDFDNELKRKTELLPQGMDBLKLVOLIAYT 243
DR	PDB3009; GDDP; 1.	Db	185	QEFGKQAVATLKLQYGTKTQKTDQFDNLKRVKTDLIPKGWDVKLVOLIAYT 244
KW	Complete proteome.	Qy	244	DWKETOKDPKGYMWNYYNDMFKPGAMADEVVYKADGVGPWMVLVNKERSKPDNTIVYTP 303
SEQUENCE	358 AA; 41162 MW; 9BEF514318C30413 CRC64;	Db	245	DWHETBKNAQGKWNYYDWMFKPGAMADEVVYKADGVGPWMVLDDKNSKAGDDKYTP 304
Matches 297;	Conservative	Qy	304	LVEKLQAYNVEVHPTVRKDALPEFTDYNQMDALINKSGATGVTFDPDTGVFL 360
Best Local Similarity	82.3%; Pred. No. 2.e-101;	Db	305	MVADTAKTXKMBLHPPTVRKDALPAFPFTDYNQMDALYNHAGTGLFTDPDTGVFL 361
Matches 297;	Mismatches 31; Indels 8	Qy	RESULT 4	
1	MDPQTIALSLIAAGVLAGCSSHSSNNMANTQMSDKIIIAHRSAGSLPEHTLESKALAFKA 60	QBRIT6	PRELIMINARY;	PRT; 357 AA.
Db	1 MKLKTL-VAIALLSLTACSS-SMMKN----DDKLIIIAHRSAGSLPEHTLESKALAFG 54	AC	QBRIT6;	PRT; 357 AA.
Qy	61 QDADYEQDLAMTKDQGRLVYIHDHFLDGLTDYAKKPFPHRKDGRTYYTDFTLKEIOSL 120	DT	01-JUN-2002 (TREMBrel. 21, Created)	
Db	55 QDADYEQDLAMTKDQRLVYIHDHFLDGLTDYAKKPFPHRKDGRTYYTDFTLKEIOSL 114	DT	01-JUN-2002 (TREMBrel. 21, Last sequence update)	
Qy	121 MTENFETKDGQAOQVNPRLWKSHPHRHTPEDELEPIQLEKSTGKCGYIYPEIKAPW 180	DT	01-JUN-2003 (TREMBrel. 24, Last annotation update)	
Db	115 MTENFETKDGQAOQVNPRLWKSHPHRHTPEDELEPIQLEKSTGKCGYIYPEIKAPW 174	DT	Glycerocephoryl diester phosphodiesterase (EC 3.1.4.46).	
Qy	181 FHHONGKDIAAETLKLKQYGDRTKTDQMDVLYQDFDNELKRKTELLPQGMDBLKLVOLI 240	GN	OrderedLocusNames=PN1908;	
Db	175 LHHKEGKDIAAETLKLKQYGDRTKTDQMDVLYQDFDNELKRKTELLPQGMDBLKLVOLI 234	OS	Fusobacterium nucleatum (subsp. nucleatum).	
Qy	241 AYTDWKETOKDPKGYMWNYYNDMFKPGAMADEVVYKADGVGPWMVLVNKERSKPDNTV 300	OC	Bacteria; Fusobacteria; Fusobacteriaceae;	
Db	235 AYTDWKETOKDPKGYMWNYYDWMFKPGAMADEVVYKADGVGPWMVLVNKERSKPDNTV 294	OC	Fusobacterium nucleatum	
Qy	301 YTPLVKELQAYNVEVHPTVRKDALPEFTDYNQMDALINKSGATGVTFDPDTGVFL 360	NCBI_TaxID=76856;		
Db	295 YTPLVKELQAYNVEVHPTVRKDALPEFTDYNQMDALINKSGATGVTFDPDTGVFL 354	RN	[1]	
Qy	361 K 361	RP	SEQUENCE FROM N.A.	
Db	355 K 355	RC	STRAIN=ATCC 25586;	
RESULT 3	Q65R11	PRELIMINARY;	MEDLINE=21886394; PubMed=116899109;	
AC	Q65R11;	PRT; 368 AA.	DOI=10.1128/JB.18.7.2005-2018.2002;	
DB	25-OCT-2004 (TREMBrel. 28, Created)	RA	Kapral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidas A.,	
DT	25-OCT-2004 (TREMBrel. 28, Last sequence update)	RA	Bhattacharya A., Bartram R., Gardner W., Grechkin G., Zhu J.,	
DT	25-OCT-2004 (TREMBrel. 28, Last annotation update)	RA	Vasileva O., Chu L., Kogan Y., Chaga O., Goitsman B., Bernai A.,	
DR	UgpQ protein	RA	Larsen N., Di Souza M., Walunas T., Pusch G., Haselkorn R.,	
DE	Name=ugpQ; ORFNames=MS1992;	RA	Ponstein M., Kyriides N.C., Overbeek R.;	
GN	Manneheimia succiniciproducens MBE155B.	RT	"Genome sequence and analysis of the oral bacterium Fusobacterium nucleatum strain ATCC 25586";	
OS	Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;	RT	J. Bacteriol. 184:2005-2018(2002).	
OC	Pasteurellaceae; Mannheimia.	RI	DR EMBL; AB010493; AU94007.1; -;	
NCBI_TaxID=221988;	OX	DR GO; 0008889; P: glycerophosphodiester phosphodiesterase act. . . ; IEA.		
RN	[1]	DR GO; GO:0006071; P: glycerol metabolism; IEA.		
RP	SEQUENCE FROM N.A.	DR InterPro; IPR04139; GDPD; 1.		
RC	STRAIN=MBSL55;	DR Pfam; PF03009; GDPD; 1.		
RA	Hong S.H., Kim J.S., Lee S.Y., In Y.H., Choi S.S., Rih J.-K.,	KW Complete Proteome; SEQNCB 357 AA; 41111 MW; BBBCF3F3D3PDS3B0 CRC64;		
RA	Kim C.H., Jeong H., Hur C.-G., Kim J.G., Kim J.-G.,	Query Match 60.5%; Score 1498.5%; DB 2; Length 357;		
RA	"The genome sequence of the capnophilic rumen bacterium Mannheimia succiniciproducens"; Nat. Biotechnol. 0.0 (2004);	Best Local Similarity 77.7%; Pred. No. 6.2e-96;		
RT	EMBL; AE016822; AAU38599.1; -;	Matches 278; Conservative 35; Mismatches 42; Indels 8 3; Gaps 1;		
RL	SEQUENCE 368 AA; 42139 MW; 9E0A0D60B029873C CRC64;	Qy	8 LSLLAAGVLAGCSSHSSNNMANTQMSDKIIIAHRSAGSLPEHTLESKALAFQA 63	
DR	Query Match 60.6%; Score 1501.5%; DB 2; Length 368;	Db	68 QDLAMTKDQGRLVYIHDHFLDGLTDYAKKPFPHRKDGRTYYTDFTLKEIOSLEMENFET 127	
Matche	Best Local Similarity 77.0%; Pred. No. 4e-96;	Qy	58 QDLAMTKDQGRLVYIHDHFLDGLTDYAKKPFPHRKDGRTYYTDFTLKEIOSLEMENFET 117	
279;	Mismatches 37; Indels 44; Gaps 1;	Db	128 KDGKQAOQVNPRLWKSHPHRHTPEDELEPIQLEKSTGKCGYIYPEIKAPWKHQNKG 187	
SEQUENCE	368 AA; 42139 MW; 9E0A0D60B029873C CRC64;	Db	118 KDGKQAOQVNPRLWKSHPHRHTPEDELEPIQLEKSTGKCGYIYPEIKAPWKHQNKG 177	
Qy	4 KTLAISLLAAGVLAGCSSHSSNNMANTQMSDKIIIAHRSAGSLPEHTLESKALAFQA 63	Qy	188 DIAATLAKKYLKQYDQKCTDMVATLQTFDENELKRKTELLPQMDLKVOLIAYTDWK 247	
6 KTLAISLLAAGVLAGCSSHSSNNMANTQMSDKIIIAHRSAGSLPEHTLESKALAFQA 64				

RESULT 5

Q7P6W6 PRELIMINARY; PRT; 309 AA.

ID Q7P6W6 PRELIMINARY; PRT; 309 AA.

AC Q7P6W6; PRELIMINARY; PRT; 309 AA.

DT 01-MAR-2004 (TREMBrel. 26, Last sequence update)

DT 01-MAR-2004 (TREMBrel. 26, Last annotation update)

DT 01-MAR-2004 (TREMBrel. 26, Last annotation update)

DE Glycerophosphodiester phosphodiesterase (EC 3.1.4.46).

GN Name=PNV198; Order=locusName=ECB124, Z3497;

OS Escherichia coli O157:H7.

OC Bacterium; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=83134;

RN [1]

SEQUENCE FROM N.A.

RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;

RX MEDLINE=1074935; PubMed=11206551; DOI=10.1038/35054089;

RA Perera N.T., Blunkett G.III, Burland V., Maiu B., Glaeser J.D., Rose D.J., Maynew G.F., Evans P.S., Boutin J., Kirkpatrick H.A., Postai G., Hackett J., Klink S., Boutin A., Gregor J., Miller L., Grobeck B.J., Davis N.W., Lim A., Dimalanta E.T., Apoaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C., Welch R.A., Blattner F.R.;

RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7." Nature 409:529-533 (2001).

RN [2]

SEQUENCE FROM N.A.

RC STRAIN=O157:H7 / RIMD 0509932 / EHEC;

RX MEDLINE=1156321; PubMed=11258796;

RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K., Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T., Iida T., Takami H., Honda T., Sasaki N., Yasunaga T., Kubara S., Shiba T., Hattori M., Shinagawa H.;

RT "Complete genome sequence of enterohaemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12."

RL DNA Res. 8:11-22 (2001).

DR EMBL; AE05456; AACG57370.1; -.

DR EMBL; AP00256; BAB36547.1; -.

DR PIR; D91019; DR91019.

DR PIR; P85863; P85863.

DR GO; GO:0006078; P: glycerophosphodiester phosphodiesterase act. . . IEA.

DR GO; GO:0006079; P: glycerol metabolism; IEA.

DR InterPro; IPR004129; GPD.

DR Pfam; PF03009; GDPD; 1.

DR PIR; D91019; DR91019.

DR PIR; P85863; P85863.

DR GO; GO:0006078; P: glycerophosphodiester phosphodiesterase act. . . IEA.

DR InterPro; IPR004129; GPD.

KW Hydrolase.

SEQUENCE 309 AA; 36073 MW; CE2F4189TA63A8D7 CRC64;

Query Match 55.0%; Score 1363; DB 2; Length 309;

Best Local Similarity 80.4%; Pred. No. 1.3e-86;

Matches 246; Conservative 30; Mismatches 30; Indels 0; Gaps 0;

RESULT 6

Q8XK17 PRELIMINARY; PRT; 358 AA.

ID Q8XK17 QAC3 ;

AC Q8XK17 ;

DT 01-MAR-2002 (TREMBrel. 20, Last sequence update)

DT 01-MAR-2002 (TREMBrel. 20, Last annotation update)

DT 25-OCT-2004 (TREMBrel. 28, Last annotation update)

DE Glycerocephosphodiester phosphodiesterase, periplasmic.

GN Name=g190; Order=locusName=ECB124, Z3497;

OS Escherichia coli O157:H7.

OC Bacterium; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=83134;

RN [1]

SEQUENCE FROM N.A.

RC STRAIN=O157:H7 / RIMD 0509932 / EHEC;

RX MEDLINE=1156321; PubMed=11258796;

RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K., Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T., Iida T., Takami H., Honda T., Sasaki N., Yasunaga T., Kubara S., Shiba T., Hattori M., Shinagawa H.;

RT "Complete genome sequence of enterohaemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12."

RL DNA Res. 8:11-22 (2001).

DR EMBL; AE05456; AACG57370.1; -.

DR EMBL; AP00256; BAB36547.1; -.

DR PIR; D91019; DR91019.

DR PIR; P85863; P85863.

DR GO; GO:0006078; P: glycerophosphodiester phosphodiesterase act. . . IEA.

DR GO; GO:0006079; P: glycerol metabolism; IEA.

DR InterPro; IPR004129; GPD.

DR Pfam; PF03009; GDPD; 1.

DR Complete proteome.

SEQUENCE 358 AA; 40856 MW; OCAB969376591EF CRC64;

Query Match 50.9%; Score 1261.5; DB 2; Length 358;

Best Local Similarity 64.7%; Pred. No. 1.8e-79;

Matches 231; Conservative 55; Mismatches 64; Indels 7; Gaps 2;

QY 4 KTLAISLLAAGVLAGCCSSHSNNMANTOMKSDKLIIAHRGASGVLPEHTLESKALAAQQA 63

DB 6 KNLSMAMMSTIVNSAAAAD----SNEKCVIUAHRSASGVLPEHTLPAKAMAYAQQA 59

QY 64 DYLEBDLAMTKDGRUVVHDHFGLTDVAKKEPHHRKGDKRYVITDFTLKEIOSLEMTE 123

DB 60 DYLEBDLAMTKDGRUVVHDHFGLTDVAKKEPHHRKGDKRYVITDFTLKEIOSLEMTE 119

QY 124 NFETKDGKQDQAQVYPNRPFLWKSHERIHTPDETBIFQGLEKSTGRKVGYPEIKAPWFHH 183

DB 120 GFDIENGKKVQTYPORFPNGKSDPFRHTFEEIEVQGLNHSIGKNGIYPEIKAPWFHH 179

QY 184 QNGKIAAETLKVKKYGDKTDMYQLOTFDNELKRKIKTELLPMGMQLKLVQL 243

DB 180 QEGKDIAAKTLLEVLRKKGTYGTGKODRVTLOCDFDALKNLKRNELPKGMNDLNLYQLIATT 239

QY 244 DWKETQEDPKGTVWNNYDWMFKPGAMAEVVKAADGVGPWMVLVNKEESKPDDNI 303

DB 240 DNNETQQKQPDGSWVNNYDWMFKPGAMAEVVKAADGVGPWMVLVNKEESKPDDNI 298

QY 300 VTPVYKELAQVNVBHPVTPVRKDALPEFTDVONYMDALLINKSATGVTDFPTGVEF 359

DB 302 LENQRN 307

QY 360 LKGKIKS 3 65

DB 304 LVKBLAQVNVBHPVTPVRKDALPEFTDVONYMDALLINKSATGVTDFPTGVEFPL 360

RESULT 7									
Q8CVV6	Q8CVV6	PRELIMINARY;	PRT;	358 AA.					
QAC	QAC								
DTT	DTT	23, Created)							
DTT	DTT	(TrEMBLrel. 23, Last sequence update)							
DTT	DTT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)							
DTT	DTT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)							
DE	DE	Glycerophosphoryl diester phosphodiesterase, periplasmic (EC 3.1.4.46).							
GEN	GEN	Name=gipQ; OrderedLocusNames=c2780;							
OS	OS	Bacillus coli O6.							
OC	OC	Bacterium; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Bacteroidales; Escherichia.							
NCBI_TaxID	NCBI_TaxID	NCBI_TaxID=217992;							
[1]		SEQUENCE FROM N.A.							
RC	RC	STRAIN=06_H1 / CFT073 / ATCC 700928;							
RX	RX	MDLINE=22388234; PubMed=11471157; ID=10.1107/pnas.252529799;							
RA	RA	Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P., Rosko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D., Mayhew G.P., Rose D.J., Zhou S., Schwartz D.C., Perna N.T., Mobley H.L.T., Donnenberg M.S., Blattner F.R.;							
RA	RA	"extensive mosaic structure revealed by the complete genome sequence of uropathogenic Escherichia coli.";							
RT	RT	Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).							
RL	RL	EMBL: AB016563; AACN81234.1; -							
DR	DR	GO:0010089; P: glycerophosphodiester phosphodiesterase act. . . ; IEA.							
DR	DR	GO: GO_0016787; P: hydrolase activity; IEA.							
DR	DR	GO: GO_0006071; P: glycerol metabolism; IFA.							
DR	DR	InterPro: IPR004129; GPD.							
DR	DR	PFam: PF03009; GBDP; 1.							
KW	KW	Complete proteome; Hydrolase.							
SEQ	SEQ	SEQUENCE 358 AA; 40826 MW; C633D35SEA6591F8 CRC64;							
Query Match		50.6%; Score 1254.5; DB 2; Length 358;							
Best Local Similarity		64.4%; Pred. No. 5-6-79;							
Matches	230;	Conservative 55; Mismatches 65; Indels 7; Gaps							
DY	DY	4 KTLAISLIAAGVLAGCSSHSSPANTOMKSKKIIAHRGASSYLPENTLBSVALAPAAQA 6							
DB	DB	6 KNLSMADMSTMSTIMGSSAMAAD --- -SNEKIVIAHRGASSYLPENTLPAKAMAYAQGA 5							
DY	DY	64 DYLEQDLAMTKDGRVYTHDHFGLDTVAKKFPHRDKGYYVDFTLKEDIOSLEMT 1							
DB	DB	60 DYEQDLWMTKDDHLWVHDHLYDRVTDAEFDPDRDKDGYYALDFTLBIIKS1KFTE 1							
DY	DY	124 NFETKDQKAOAQYPNRPFLWKSHFRINTFEDIEPIOLEKSTGKVGYIPEKIKPWFH 1							
DB	DB	120 GFDIENGCKVOTYPRGFPMGKSDFRVHTEEBEETFYQLNHSSTGKNIGIYIPEKIKPWFH 1							
DY	DY	184 QNGKDIAMETLKVYKCYGYDKEITDMYLQTFFENELKRITKELPONGMDJRLVQIAYT 2							
DB	DB	180 QEGKDIAKKTLEVYKCYGYDKEITDMYLQTFFENELKRITKELPONGMDJRLVQIAYT 2							
DY	DY	244 DNKETQEKDPKGKQYWNTNYDMMFPGKAMAEVYKYADVGPGSKWMLYNKEBSKPDNIVYTP 3							
DB	DB	240 DNNETOQQPDPGSWNTNYDNMFPGKAMQVPEYADGIGDPYHMLI-EETSQGNIKLTG 2							

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage for commercial enterprises requires a license agreement (See <http://www.isb-sib.ch/announce>).

— 1 —

Qy	4	KTLALSLLAAGVLAGCSSHSSNMANTQMSDKLIIAHRGASGLPLPEHTLESKALAPAQA	63	Db	63	YLEDQDLVLTQNAlAIVLHDHYLDRTDVAERFQRARKDGRPYAIDFTLKEIKSLKPTEG	122		
Db	25	KNLSMAMTMSITMGSSAMAAD-----SNEKLVIAHRGASGLPLPEHTLPARAMAYAQAA	78	Qy	125	FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHQ	184		
Db	64	DYLEQDLAMTKDGRVLVTHDHELDGLTDVAKEPFHHRKGRRYVIFDTLKEIQSLEMTE	123	Db	123	FPIKGKQVQSYPNRFPMWKSDPRHTPOEEBFIQCLNHSTGKNIGLYPEIKAPWPWHQ	182		
Qy	79	DYLEQDLAMTKDHLVYLDRTDVADEPDRARKDGRYYAIDFTLDBIKSLKFT	138	Qy	185	NGKDIAREBTLKVLYKQYGDKEKTDMMYLQTFDFNELKRKTELLPOMGMDLKLVQLIAYTD	244		
Db	124	NFETKGKQAOQYYPNRFPFLWKSFRITFEDETEFIOGLEKSTGKVGYPIKAQWPFFH	183	Db	183	BGKDIAREBTLTVLQGYTEKSDRVFLQSDVAELKRKTELOPQMKMDLKLVQLIAYTD	242		
Qy	139	GFDIENGKVKVQYPPGRPMGKSDPRVHTPEEEFVQLNNSTGKNIGYPIKAQWPFFH	198	Qy	245	WKETQBKDPKGWVNTNNYDNMKPGAMAEVYKADGVPGMMLVNKEESKPDNVITYTPL	304		
Qy	184	ONGKDIAREBTLKVLYKQYGDKEKTDMMYLQTFDFNELKRKTELLPOMGMDLKLVQLIAYT	243	Db	243	WHTYERQADSGSWNTNNYDNMKPGAMKQIATYDGGPDYHMLVOK-DSTPDSIIFTNMM	301		
Db	199	QEBKDIAMKTLEBLVLYKQYGDKEKTDMMYLQTFDFNELKRKTELLPOMGMDLKLVQLIAYT	258	Qy	305	VKBALQINVEWHPYTRKDALPEFTDVNQMTDALINKSGATGVFDPPDTGVEPLK	361		
Qy	244	DWKEHQDPKPGWVNTNNYDNMKPGAMAEVYKADGVPGMMLVNKEESKPDNVITYTP	303	Db	302	VKEAHQHKLQVHPFTARADRUPYAKDIDOLYDILTHQADVGLFTDPPDKAVDPLK	358		
Db	259	DWNETQKQPDGSWWVNTNNYDNMKPGAMQVAEYADGVPDMI-BETSGPNIKLTG	317	RESULT 12					
Qy	304	LVRBLAQINVEWHPYTRKDALPEFTDVNQMTDALINKSGATGVFDPPDTGVEPLK	360	Q6ZNG7	PRELIMINARY;	PRT;	356 AA.		
Db	318	MVQDAQQMKLUVHPYTVRSKDPVPEYTPDVNLQDYALYNKAGYNGLFDPPDKAVKFL	374	ID	Q8ZNG7	PRELIMINARY;	PRT;	356 AA.	
RESULT 11						AC	RC		
Q6CZI4		PRELIMINARY;	PRT;	361 AA.		DT	DT		
AC	Q6CZI4;					01-MAR-2002	(TREMBLrel. 20, Created)		
DT	25-OCT-2004	(TREMBLrel.	28	Last sequence update)		DT	01-MAR-2003		
DT	25-OCT-2004	(TREMBLrel.	28	Last annotation update)		DT	(TREMBLrel. 24, Last annotation update)		
DB	*	Glycerophosphoryl diester phosphodiesterase, periplasmic	(EC			DE	Periplasmic glycerophosphodiester phosphodiesterase (EC		
DE	3.1.4.46).					DE	Name=glpO; OrderedLocusNames=STM2282;		
GN	*	Name=glpO; OrderedLocusNames=EC4167;				GN	Name=glpO; OrderedLocusNames=STM2282;		
OC	Erwinia carotovora (subsp. atroseptica) (Pectobacterium atrosepticum).					OS	Salmonella typhimurium.		
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;					OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;		
NCBI_TaxID	29471;					NCBI_TaxID	Salmonella.		
RN	[1]					RN	SEQUENCE FROM N.A.		
RP	SEQUENCE FROM N.A./	ATCC BAA-672;				RP	SEQUENCE FROM N.A.		
RC	PubMed=15233089; DOI=10.1073/pnas.0402424101;					RC	Strain=Lt2;		
RA	Bell K.S., Sebaiha M., Pritchard L., Holden M.T.G., Hyman L.J.,					RA	Medline=21534948; PubMed=11677609; DOI=10.1073/pnas.0402424101;		
RA	Holewko M.C., Thomson N.R., Bentley S.D., Church L.J.C., Mungall K.,					RA	McClelland M., Sanderson K.E., Spielet J., Clifton S.W., Latreille P.,		
RA	Atkin R., Bason N., Brooks J., Chillingworth T., Clark K., Doggett J.,					RA	Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,		
RA	Fraser A., Hance Z., Hausrath H., Jagels K., Moule S., Norbertczak H.,					RA	Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,		
RA	Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,					RA	Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M., Waterston R., Wilson R.K.;		
RA	Salmond G.P.C., Birch P.R.J., Parkhill J., Toth I.K.,					RA	"Complete genome sequence of <i>Salmonella enterica</i> serovar Typhimurium		
RT	"Genome sequence of the enterobacterial phytopathogen <i>Erwinia</i>					RT	RT LT2";		
RT	carotovora	subssp. atroseptica and characterization of virulence				RL	Nature 413:852-856 (2001).		
RT	factors."					RA	EMBL; AE008822; AA21831;		
RL	Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110 (2004).					RA	GO; GO:0008889; F:glycerophosphodiester phosphodiesterase act. . . ; IEA.		
EMBL	BX90851; CAG77064.1;					RA	GO; GO:0016787; F:hydrolyase activity; IEA.		
DR	GO; GO:0008889; F:glycerophosphodiester phosphodiesterase act. . . ; IEA.					RA	DR; GO:0006071; P:glycerol metabolism; IEA.		
DR	InterPro; IPR004129; GDDP; 1.					RA	Pfam; PF03009; GDDP; 1.		
DR	Complete Proteome; Hydrolase;					RA	Complete proteome; Hydrolase.		
SQ	SEQUENCE 361 AA; 41551 MW; EB3B20B4555ABS56 CRC64;					SQ	SEQUENCE 356 AA; 40425 MW; 527261A667E09C44 CRC64;		
Query Match						Qy	Score 1235; DB 2;		
Best Local Similarity						Qy	Length 356;		
Matches 231; Conservative						Qy	Best Local Similarity 64.5%; Pred. No. 1.3e-77; Mismatches 47; Mismatches 231; Conservative 47; Mismatches 70; Indels 10; Gaps 2;		
Matches 228; Conservative						Db	6 KNLSVAlMLAGMTTSGA-----VAEKKVLAHRGASGLPHTLPKAMVQAQGA 56		
Matches 228; Conservative						Db	4 KTLALSLLAAGVLAGCSSHSSNMANTQMSDKLIIAHRGASGLPLPEHTLESKALAPAQA 63		
Matches 228; Conservative						Db	5 DYLEQDLVLTQNAlAIVLHDHYLDRTDVAERFQRARKDGRPYAIDFTLKEIKSLKPTEG 122		
Matches 228; Conservative						Db	57 DYLEQDLVLTQNAlAIVLHDHYLDRTDVAERFQRARKDGRPYAIDFTLKEIKSLKPTEG 122		
Matches 228; Conservative						Db	6 KNLSVAlMLAGMTTSGA-----VAEKKVLAHRGASGLPHTLPKAMVQAQGA 56		
Matches 228; Conservative						Db	64 DYLEQDLAMTKDGRVLVTHDHELDGLTDVAKEPFHHRKGRRYVIFDTLKEIQSLEMTE		
Matches 228; Conservative						Db	117 GFDIENGKVKQTYPGRPMGKSDPRHTFBBTPEVQGLNHSSTGKNIGYPIKAQWPWHH 176		
Matches 228; Conservative						Db	117 GFDIENGKVKQTYPGRPMGKSDPRHTFBBTPEVQGLNHSSTGKNIGYPIKAQWPWHH 176		
Matches 228; Conservative						Db	124 NFETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 183		
Matches 228; Conservative						Db	125 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 184		
Matches 228; Conservative						Db	126 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 185		
Matches 228; Conservative						Db	127 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 186		
Matches 228; Conservative						Db	128 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 187		
Matches 228; Conservative						Db	129 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 188		
Matches 228; Conservative						Db	130 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 189		
Matches 228; Conservative						Db	131 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 190		
Matches 228; Conservative						Db	132 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 191		
Matches 228; Conservative						Db	133 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 192		
Matches 228; Conservative						Db	134 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 193		
Matches 228; Conservative						Db	135 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 194		
Matches 228; Conservative						Db	136 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 195		
Matches 228; Conservative						Db	137 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 196		
Matches 228; Conservative						Db	138 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 197		
Matches 228; Conservative						Db	139 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 198		
Matches 228; Conservative						Db	140 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 199		
Matches 228; Conservative						Db	141 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 200		
Matches 228; Conservative						Db	142 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 201		
Matches 228; Conservative						Db	143 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 202		
Matches 228; Conservative						Db	144 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 203		
Matches 228; Conservative						Db	145 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 204		
Matches 228; Conservative						Db	146 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 205		
Matches 228; Conservative						Db	147 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 206		
Matches 228; Conservative						Db	148 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 207		
Matches 228; Conservative						Db	149 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 208		
Matches 228; Conservative						Db	150 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 209		
Matches 228; Conservative						Db	151 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 210		
Matches 228; Conservative						Db	152 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 211		
Matches 228; Conservative						Db	153 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 212		
Matches 228; Conservative						Db	154 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 213		
Matches 228; Conservative						Db	155 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 214		
Matches 228; Conservative						Db	156 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 215		
Matches 228; Conservative						Db	157 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 216		
Matches 228; Conservative						Db	158 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 217		
Matches 228; Conservative						Db	159 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 218		
Matches 228; Conservative						Db	160 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 219		
Matches 228; Conservative						Db	161 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 220		
Matches 228; Conservative						Db	162 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 221		
Matches 228; Conservative						Db	163 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 222		
Matches 228; Conservative						Db	164 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 223		
Matches 228; Conservative						Db	165 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 224		
Matches 228; Conservative						Db	166 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 225		
Matches 228; Conservative						Db	167 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 226		
Matches 228; Conservative						Db	168 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 227		
Matches 228; Conservative						Db	169 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 228		
Matches 228; Conservative						Db	170 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 229		
Matches 228; Conservative						Db	171 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 230		
Matches 228; Conservative						Db	172 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 231		
Matches 228; Conservative						Db	173 FETKGKQAOQYYPNRFPFLWKSFRHIFTFEDTEBFIOGLEKSTGKVGYPIKAQWPWHH 232		
Matches 228; Conservative						Db	174 FETKGKQAOQ		

Qy	244 DWKETQDKPKGYWVNNYDWMFKPGAMAEVVKYADGVGPWGMLVNKEESKPDNIVTP 303	Qy	184 QNGKDIQAETFLKUKKKYGYDKKCTDMVYLOTFDNELKRKTTELLQMGMDKLKVLIAYT 243
Db	237 DNNETQQKQPKGRWVNNYDWMFKGAMKVAEYDGIGDHYMLV-AEGSTKGNIKLJG 295	Db	177 QEGKDIQAETAKLEVTKKYGIGKQNNVYLCQFDVAELKRKTNELQMGMDKLKVLIAYT 236
Qy	304 LVKELAQVNEVHPYTVRKDALPEFTDVQMYDALLNKSATGVFTDPDTGVEFLK 361	Qy	244 DWKETQEKDPKGYWVNNYDWMFKPGAMAEVVKYADGVGPWGMLVNKEESKPDNIVTP 303
Db	296 MVQDAHQNKVVHPPTVRADOLPDTAVNQLYDLYNKAQVDGLFTDPDKAVMFLQ 353	Db	237 DNNETQQKOPDGRWVNNYDWMFKEGAMKQVAEYDGIGDHYMLV-ARGSTKGNIKLJG 295
RESULT 13			
Q8Z556	PRELIMINARY;	PRT;	356 AA.
ID	Q8Z556; Q7CB71;		
AC	Q8Z556; Q7CB71;		
DT	01-MAR-2002 (TREMBLrel. 20, Created)	DT	07MAR07 PRELIMINARY;
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)	AC	Q7M2Y7 PRELIMINARY;
DT	25-OCT-2004 (TREMBLrel. 28, Last annotation update)	DT	01-MAR-2004 (TREMBLrel. 26, Created)
DE	3.1.4.6)	DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)
GN	Name=glpQ; OrderedLocusNames=STY2511, t0582;	DB	Glycerophosphoryl diester phosphodiesterase, periplasmic
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;	DB	(Glycerophosphoryl diester phosphodiesterase).
OC	Enterobacteriaceae; Salmonellae.	GN	Name=glpQ; OrderedLocusNames=plu120;
OX	[1] NCBI_TaxID=601;	OS	Photorhabdus luminescens (subsp. laumontii).
RN	SEQUENCE FROM N.A.	OC	Bacteria; Protobacteria; Gammaproteobacteria; Enterobacteriales;
RP	STRAIN=CT18;	OC	Enterobacteriaceae; Photorhabdus.
RX	MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;	OX	NCBI_TaxID=141679;
RA	Parkhill J., Dougan G., James K.D., Thomson N.R., Pickett D., Wain J.,	RN	[1]
RA	Barker C.M., Mungall K.L., Bentley S.D., Holden M.R., Sebastianha M.,	RP	SEQUENCE FROM N.A.
RA	Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,	RC	STRAIN=CT18;
RA	Cronin A., Davis P., Davies R.M., Dowd L., White N., Parrar J.,	RX	MEDLINE=22957627; PubMed=14528314;
RA	Feltwell T., Hamlin N., Haque A., Hien T.T., Holoyd S., Jagels K.,	RA	Dechaud E., Rusniok C., Frangeul L., Buchrieser C., Givaudan A.,
RA	Krogh A., Larsen T.S., Leather S., Monle S., O'Gaoar P., Parry C.,	RA	Taorat S., Bcgs S., Bourreux-Bude C., Chandler M., Charles J.-F.,
RA	Quail M.A., Rutherford K.M., Simmonds M.N., Skelton J., Stevens K.,	RA	Dassa E., Derose R., Deszelie S., Freyssinet G., Gaudriault S.,
RA	Whitehead S., Barrell B.C.;	RA	Medigue C., Lanois A., Powell K., Signier P., Vincent R., Wingate V.,
RT	"Complete genome sequence of a multiple drug resistant <i>Salmonella</i> RT	RA	Zouine M., Glaser P., Bemare N., Danchin A., Kunst F.;
RT	enterica serovar Typhi CT18." RT	RT	"The genome sequence of the entomopathogenic bacterium <i>Photorhabdus</i> luminescens";
RN	[2]	RT	Nature 413:848-852 (2001).
RP	SEQUENCE FROM N.A.	RL	Nat. Biotechnol. 21:1307-1313 (2003).
RC	STRAIN=TY2 / ATCC 700931;	DR	EMBL; BX571872; CAB16432.1; -.
RX	MEDLINE=22531367; PubMed=12644504;	DR	PhytoLib; plu4120; -.
RX	DOI=10.1128/JB.185.7.2330-2337.2003;	DR	GO; GO:0008889; P: glycerophosphodiester phosphodiesterase act. . . ; IEA.
DR	Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,	DR	GO; GO:0006071; P: glycerol metabolism; IEA.
RA	Burland M., Kodoyianni V., Schwartz D.C., Blattner F.R./	DR	InterPro; IPR004129; GDDP.
RA	"Comparative genomics of <i>Salmonella enterica</i> serovar Typhi strains Ty2	DR	Pfam; PF03009; GDDP.
RT	and CT18." RT	KW	Complete proteome.
RL	J. Bacteriol. 185:2330-2337 (2003).	SQ	SEQUENCE 358 AA; 4118.2 MW; OBBB12D55785ABC5 CRC64;
DR	EMBL; AL6227274; CAD07514.1; -.	Query Match	48.9%; Score 1210.5; DB 2; Length 358;
DR	EMBL; AB016836; AA088288.1; -.	Best Local Similarity	62.0%; Pred. No. 6.4e-76; Indels 11; Gaps 2;
DR	GO; GO:0008889; P: glycerophosphodiester phosphodiesterase act. . . ; IEA.	Matches 22;	Conservative 54; Mismatches 71;
DR	GO; GO:0016787; P: hydrolase activity; IEA.	Matches 231;	Score 1233; DB 2; Length 356;
DR	Pfam; PF01009; GDDP; 1.	Query Match	49.8%; Score 1233; DB 2; Length 356;
KW	Complete proteome; Hydrolase.	Best Local Similarity	64.5%; Pred. No. 1.7e-77; Indels 10; Gaps 2;
SQ	SEQUENCE 356 AA; 40437 MW; 51AF570169108CAA CRC64;	Matches 231;	Conservative 47; Mismatches 70;
Db	4 KTLAISLLAAAGVLAGCSSHSSNMANTQMSDKIIIAHRGASGYLPDEHTLESKALAPAQQA 63	Query	4 KTLAISLLAAAGVLAGCSSHSSNMANTQMSDKIIIAHRGASGYLPDEHTLESKALAPAQQA 63
Db	6 KNLSVALMLAGMTISGA -----VAAEKVVAHRGASGYLPDEHTLPKAMVYQQA 56	Db	6 DYLEQDLIATMKTGDRGIVHDFLQGLTDVAKKFPHRHKDGRYYTIDFTLKEIOSLEMTE 123
Qy	64 DYLEQDLIATMKTGDRGIVHDFLQGLTDVAKKFPHRHKDGRYYTIDFTLKEIOSLEMTE 123	Qy	184 QNGKDIQAETFLKUKKKYGYDKKCTDMVYLOTFDNELKRKTTELLQMGMDKLKVLIAYT 243
Db	57 DYLEQDLIATMKTGDRGIVHDFLQGLTDVAKKFPHRHKDGRYYTIDFTLKEIOSLEMTE 116	Db	176 KEGD13TCKVLAFLKAYG7TKKSDIYLCQFDTNELKRKTNELQPLGMDKLKVLIAYT 235
Qy	124 NFETDKGKOAQYQYPNRFLPKSHFRHTFEDEIEFGLEKSTGKVGTYPEIKAQWFFH 183	Qy	244 DWKETQEKDPKGYWVNNYDWMFKEGAMKQVAEYDGIGDHYMLV-YSTPTNIKLTN 294
Db	117 GFDIENGKVKVOTYGRFMGRSKDSDFRIFTFEEBIEFPVQGLNHSTGRNIGTYPEIKAQWFFH 176	Db	236 DNNETYEKQSDGRWVNNYDWMFKEGAMKQVAEYDGIGDHYMLV-KEV-YSTPTNIKLTN 294
Qy	304 LVKELAQVNEVHPYTVRKDALPEFTDVQMYDALLNKSATGVFTDPDTGVEFLK 361	Qy	304 LVKELAQVNEVHPYTVRKDALPEFTDVQMYDALLNKSATGVFTDPDTGVEFLK 361

Db 295 LVKEAHTNLEVHPYTFIRDQLPKYATSGDQLEFDIYNQAGVDGFVFTDPDLGVKFQ 352

RESULT 15

Q66FX7	PRELIMINARY;	PRT;	371 AA.
ID Q66FX7			
AC Q66FX7			
DT 25-OCT-2004 (TREMBrel. 28, Created)			
DT 25-OCT-2004 (TREMBrel. 28, Last sequence update)			
DE Glycerophosphoryl diester phosphodiesterase precursor (EC 3.1.4.-6).			
GN Name=gipQ; ORFNames=YPTB0208;			
OS Yersinia pseudotuberculosis IP 32953.			
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
OC Enterobacteriaceae; Yersinia.			
OX NCBI_TaxID=273123;			
RN [1]			
RP SEQUENCE FROM N A.			
RC STRAIN/PI 32953;			
RX PubMed:15358858;			
RA Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,			
RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin L.V.,			
RA Rubakker R.R., Fowler J., Hinnelsbusch B.J., Marceau M., Medigue C.,			
RA Simonet M., Chenal-Francisque V., Souza B., Dacheux D., Elliott J.M.,			
RA Derbise A., Hauser L.J., Garcia B.;			
RT "Insights into the genome evolution of <i>Yersinia pestis</i> through whole			
RT genome comparison with <i>Yersinia pseudotuberculosis</i> ";			
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004);			
DR EMBL; BX936398; CAH19448.1;			
DR GO: GO:0016787; F:hydrolase activity; IEA.			
DR InterPro; IPR004129; GDFD.			
DR • Pfam; PF03009; GDFD; 1.			
KW Hydrolase; Signal.			
PT SIGNAL 1 25 Potential 1.			
SQ • SEQUENCE 371 AA; 42395 MW; 03F96587AAF6262A CRC64;			

Query Match 48.8%; Score 1210; DB 2; Length 371;

Best Local Similarity 62.7%; Pred. No. 7.3e-76; Mismatches 75; Indels 6; Gaps 4;

Matches 227; Conservative 54; Mismatches 75;

Qy 4 KTLALSLLAAGVLAGCSSHSSNMANTOMKS ---DKIIIAHRGASGYLPHTLESKALAFKA 60	Db 6 KTLALSLLAADAAA-APAYAAKPAASAKDASAIDKVIIAHRGASGYLPENSPLAKAMAYA 64
Qy 61 QOADYLEQDLMTKDGRLVVTHDHELDGLTDVAKKPFHRKDGRTYVVIDFTLKTSQLE 120	Db 65 QCADYLEQDLMTKDGRVVLVHDHYLDRTDVAAERPDARKDGRTYIAIDFTLPPEIKSLK 124
Qy 121 MTENPE-TKDQKQAQVYPNRPLWKSHFRHTFEDELFTQGLESTGKRVGIYPBKAP 179	Db 125 FTEGFDIDKNGKCKVQSYPNRPMGKSDPFRVTPQEELFIOGLNSTGRTGTYIYPBKAP 184
Qy 180 WFFHQONGKDIAAETLKVKKYGDKTTDMYLQTFDNELKRKTELLPONGMDLKLVQL 239	Db 185 WFFHQBSKDKDISTKVKLEYKEYTKNDKTVLQSFDVNELKRKNELEPKNGMDLKLVQL 244
Qy 240 IAYTDWKETQBDPKGKCYWVNTYNDWMFKPGAMAEVVYKAIDSVGPCKMMLYNKEESKPDNI 299	Db 245 IAYTDWNETYQQLPDKRKWVNTYNDWMFKPGAMKEYAHADGIPDTHMLVVKTSACKD-1 303
Qy 300 VYTPLVKELAQINNEVHPYPTVRKDALPEFFTDVNOMYDALINKSGATGVFTDPDGVBF 359	Db 304 KLTDLVKEAHASNMVTHPFTIRADRPUYATDINQLFDIYQNQAVDGFVFTDPDGQIF 363
Qy 360 LK 361	Db 364 LQ 365

THIS PAGE BLANK (USPTO)